

# Disjointness for measurably distal group actions and applications

JOEL MOREIRA<sup>†</sup>, FLORIAN K. RICHTER<sup>‡</sup> and DONALD ROBERTSON<sup>§</sup> 

<sup>†</sup> *Warwick Mathematics Institute, University of Warwick, Coventry, UK*  
(e-mail: [joel.moreira@warwick.ac.uk](mailto:joel.moreira@warwick.ac.uk))

<sup>‡</sup> *Institute of Mathematics, École Polytechnique Fédérale de Lausanne (EPFL),  
Lausanne, Switzerland*  
(e-mail: [f.richter@epfl.ch](mailto:f.richter@epfl.ch))

<sup>§</sup> *Department of Mathematics, University of Manchester, Manchester, UK*  
(e-mail: [donald.robertson@manchester.ac.uk](mailto:donald.robertson@manchester.ac.uk))

(Received 16 May 2021 and accepted in revised form 7 February 2022)

*Abstract.* We generalize Berg’s notion of quasi-disjointness to actions of countable groups and prove that every measurably distal system is quasi-disjoint from every measure-preserving system. As a corollary, we obtain easy to check necessary and sufficient conditions for two systems to be disjoint, provided one of them is measurably distal. We also obtain a Wiener–Wintner-type theorem for countable amenable groups with distal weights and applications to weighted multiple ergodic averages and multiple recurrence.

Key words: disjointness, distal systems, Kronecker factor

2020 Mathematics Subject Classification: 37A15 (Primary); 37A30, 37A35 (Secondary)

## Contents

1	Introduction	1953
1.1	Structure of the paper	1956
2	Preliminaries	1956
2.1	Measure-preserving systems	1956
2.2	Disintegrations	1957
2.3	The Kronecker factor	1957
2.4	Joinings	1958
2.5	Distal systems	1959
3	Proof of Theorem 1.3	1959
4	Quasi-disjointness for measurably distal systems	1961
4.1	Quasi-disjointness lifts through group extensions	1961
4.2	Quasi-disjointness passes to factors	1962

4.3	Quasi-disjointness is preserved by inverse limits	1966
4.4	An example	1967
5	Proof of Theorem 1.6	1967
6	Applications to multicorrelation sequences and multiple recurrence	1969
6.1	Preliminaries on nilmanifolds	1969
6.2	Proofs of Theorems 1.7 and 1.8	1971
6.3	An example of multiple recurrence	1976
7	Some open questions	1977
	Acknowledgements	1978
	References	1978

## 1. Introduction

By a  $\mathbb{Z}$  system we mean a tuple  $\mathbf{X} = (X, T, \mu_X)$  where  $X$  is a compact, metric space,  $T$  is a continuous action of  $\mathbb{Z}$  on  $X$ , and  $\mu_X$  is a Borel probability measure on  $X$  that is  $T$  invariant. Ergodic  $\mathbb{Z}$  systems  $(X, T, \mu_X)$  and  $(Y, S, \mu_Y)$  are *disjoint* if  $\mu_X \otimes \mu_Y$  is the only probability measure on  $X \times Y$  that has  $\mu_X$  and  $\mu_Y$  as its marginals and is invariant under the diagonal action  $(T \times S)^n = T^n \times S^n$ . (Any probability measure on  $X \times Y$  with these two properties is called a *joining* of the two  $\mathbb{Z}$  systems.) The notion of disjointness, introduced in Furstenberg's seminal paper [Fur67], is an extreme form of non-isomorphism. In particular, if systems have a non-trivial factor in common, then they cannot be disjoint. Furstenberg asked whether the converse is true. Rudolph [Rud79] answered this question by producing (from his construction in the same paper of a  $\mathbb{Z}$  system with minimal self-joinings) two  $\mathbb{Z}$  systems that are not disjoint and yet share no common factor.

Perhaps motivated by Furstenberg's question, Berg [Ber71, Ber72] considered the case when one of the systems is measurably distal. Recall that a  $\mathbb{Z}$  system is *measurably distal* if it belongs to the smallest class of  $\mathbb{Z}$  systems that contains the trivial system and is closed under factors, group extensions, and inverse limits; see §2 for the definitions of these notions. That the above definition of measurably distal is equivalent to Parry's original definition [Par68] in terms of separating sieves was proved by Zimmer [Zim76a] (cf. §2.5).

To describe Berg's result, recall that the *Kronecker factor* of an ergodic  $\mathbb{Z}$  system is the largest factor of the system that is isomorphic to a rotation on a compact abelian group. Berg proved that disjointness of the Kronecker factors of two ergodic  $\mathbb{Z}$  systems is equivalent to both disjointness and the absence of a common factor when one of the systems is measurably distal. We say that systems  $\mathbf{X}$  and  $\mathbf{Y}$  are *Kronecker disjoint* if their Kronecker factors are disjoint.

**THEOREM 1.1.** (Berg [Ber71, Ber72]) *Let  $\mathbf{X}$  be an ergodic and measurably distal  $\mathbb{Z}$  system and let  $\mathbf{Y}$  be an ergodic  $\mathbb{Z}$  system. The following are equivalent:*

- (i)  $\mathbf{X}$  and  $\mathbf{Y}$  are disjoint;
- (ii)  $\mathbf{X}$  and  $\mathbf{Y}$  are Kronecker disjoint;
- (iii)  $\mathbf{X}$  and  $\mathbf{Y}$  have no non-trivial common factor.

The key ingredient in the proof of Theorem 1.1 is a weakening of the notion of disjointness that is shown to be preserved by group extensions, factors, and inverse limits.

The definition of this weakened property (called ‘quasi-disjointness’ in [Ber71]) is as follows. Given ergodic  $\mathbb{Z}$  systems  $\mathbf{X}$  and  $\mathbf{Y}$ , let  $\alpha$  and  $\beta$  be the factor maps from  $\mathbf{X}$  and  $\mathbf{Y}$ , respectively, to their maximal common Kronecker factor  $K(\mathbf{X}, \mathbf{Y})$ . The  $\mathbb{Z}$  systems  $\mathbf{X}$  and  $\mathbf{Y}$  are quasi-disjoint if the property

(BQD) for almost every  $k$  in  $K(\mathbf{X}, \mathbf{Y})$  there is exactly one joining of the systems  $\mathbf{X}$  and  $\mathbf{Y}$  giving full measure to  $\gamma^{-1}(k)$

holds, where  $\gamma(x, y) = \alpha(x) - \beta(y)$ . The main results in [Ber71, Ber72] imply that  $\mathbf{X}$  and  $\mathbf{Y}$  satisfy (BQD) whenever  $\mathbf{X}$  is an ergodic and measurably distal  $\mathbb{Z}$  system and  $\mathbf{Y}$  is an ergodic  $\mathbb{Z}$  system.

In this paper, we introduce a new definition of quasi-disjointness that applies to measure-preserving actions of any countable group  $G$ . To describe it, we recall the following notions. A  $G$  system is a tuple  $(X, T, \mu_X)$  where  $X$  is a compact, metric space,  $T$  is a continuous left action of  $G$  on  $X$ , and  $\mu_X$  is a Borel probability measure on  $X$  that is  $T$  invariant. The Kronecker factor of a  $G$  system  $\mathbf{X}$  is the factor  $K\mathbf{X}$  corresponding to the subspace of  $L^2(X, \mu_X)$  spanned by functions  $f$  with the property that  $\{f \circ T^g : g \in G\}$  has compact closure. Disjointness of  $G$  systems is defined just as for  $\mathbb{Z}$  systems.

Although it is the case for ergodic  $\mathbb{Z}$  systems, the Kronecker factor of an ergodic  $G$  system cannot generally be modeled by a rotation on a compact, abelian group. Thus, it is not clear how to modify (BQD) or Berg’s proofs to apply to actions of more general groups. We instead make the following definition, which is more general and easier to handle than (BQD).

*Definition 1.2.* Two  $G$  systems  $\mathbf{X} = (X, T, \mu_X)$  and  $\mathbf{Y} = (Y, S, \mu_Y)$  are quasi-disjoint if the only joining of  $\mathbf{X}$  and  $\mathbf{Y}$  that projects to the product measure on the product  $K\mathbf{X} \times K\mathbf{Y}$  of their Kronecker factors is the trivial joining  $\mu_X \otimes \mu_Y$ .

Our first result (proved in §3) is that ergodic  $\mathbb{Z}$  systems  $\mathbf{X}$  and  $\mathbf{Z}$  are quasi-disjoint according to Definition 1.2 if and only if they satisfy (BQD), justifying the use of the terminology ‘quasi-disjoint’.

**THEOREM 1.3.** *Ergodic  $\mathbb{Z}$  systems  $\mathbf{X}$  and  $\mathbf{Y}$  are quasi-disjoint if and only if they satisfy (BQD).*

Our main result is an extension of Berg’s main results in [Ber71, Ber72] to  $G$  systems. Recall that a  $G$  system is *measurably distal* if it belongs to the smallest class of  $G$  systems that is closed under factors, group extensions, and inverse limits: notions that are defined in §2.

**THEOREM 1.4.** *If  $G$  is a countable group and  $\mathbf{X}$  is a measurably distal  $G$  system, then  $\mathbf{X}$  is quasi-disjoint from any other  $G$  system  $\mathbf{Y}$ .*

As a consequence of Theorem 1.4 we obtain the following characterizations of disjointness from a measurably distal system.

**COROLLARY 1.5.** *If  $G$  is a countable group,  $\mathbf{X}$  is a measurably distal  $G$  system, and  $\mathbf{Y}$  is a  $G$  system, then the following are equivalent:*

- (i)  $\mathbf{X}$  and  $\mathbf{Y}$  are disjoint;

- (ii)  $X$  and  $Y$  are Kronecker disjoint;  
 (iii)  $X$  and  $Y$  have no non-trivial common factor.  
 If, in addition,  $X$  and  $Y$  are ergodic, then (i)–(iii) are also equivalent to:  
 (iv) the product system  $X \times Y$  is ergodic.

If  $Y$  is measurably distal and both  $X$  and  $Y$  are ergodic, then, as pointed out to us by Glasner, the structure theory of measurably distal systems together with [Gla03, Theorem 3.30] provide an alternative approach to proving Corollary 1.5.

We offer two applications of our results. The first is a Wiener–Wintner-type result with distal weights for measure-preserving actions of countable amenable groups, proved in §5. Recall that, when dealing with actions of amenable groups, one uses Følner sequences to average orbits, where a Følner sequence in a countable group  $G$  is a sequence  $N \mapsto \Phi_N$  of finite, non-empty subsets of  $G$  such that

$$\frac{|\Phi_N \cap g^{-1}\Phi_N|}{|\Phi_N|} \rightarrow 1$$

for all  $g$  in  $G$ . Lindenstrauss [Lin01] proved that the pointwise ergodic theorem holds for actions of amenable groups along tempered Følner sequences: those for which there is  $C > 0$  with

$$\left| \bigcup_{K < N} \Phi_K^{-1}\Phi_N \right| \leq C|\Phi_N|$$

for all  $N \geq 2$ .

**THEOREM 1.6.** *Let  $G$  be a countable discrete amenable group, let  $\Phi$  be a tempered Følner sequence on  $G$  and let  $Y = (Y, S, \mu_Y)$  be an ergodic  $G$  system. For every  $\phi$  in  $L^1(Y, \mu_Y)$ , there is a conull set  $Y' \subset Y$  with the following property: for any uniquely ergodic topological  $G$  system  $(X, T)$ , with unique invariant measure  $\mu_X$ , such that the  $G$  system  $X = (X, T, \mu_X)$  is measurably distal and Kronecker disjoint from  $Y$ , and for any  $f \in C(X)$ , any  $x \in X$  and any  $y \in Y'$  we have*

$$\lim_{N \rightarrow \infty} \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} f(T^g x) \phi(S^g y) = \int f d\mu_X \int \phi d\mu_Y. \quad (1)$$

One can quickly derive the classical Wiener–Wintner theorem [WW41] from Theorem 1.6, which we do in §5.

The second type of application that we offer is to the theory of multiple recurrence. It is somewhat surprising that only Kronecker disjointness is needed for the following theorems, even though multicorrelations are typically governed by nilrotations (cf. [HK05, Lei10, Zie07]), which, in general, are of a higher complexity than rotations on compact abelian groups. However, as nilsystems are distal, Theorem 1.1 allows us to deduce disjointness of nilsystems from disjointness of their Kronecker factors.

**THEOREM 1.7.** *Let  $X = (X, T, \mu_X)$  and  $Y = (Y, S, \mu_Y)$  be ergodic  $\mathbb{Z}$  systems and assume  $X$  and  $Y$  are Kronecker disjoint. Then for every  $k, \ell \in \mathbb{N}$ , any  $f_1, \dots, f_k \in$*

$L^\infty(X, \mu_X)$  and any  $g_1, \dots, g_\ell \in L^\infty(Y, \mu_Y)$  we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k \prod_{j=1}^\ell T^{in} f_i S^{jn} g_j = \left( \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k T^{in} f_i \right) \left( \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \prod_{j=1}^\ell S^{jn} g_j \right)$$

in  $L^2(X \times Y, \mu_X \otimes \mu_Y)$ .

**THEOREM 1.8.** *Let  $(Y, S)$  be a topological  $\mathbb{Z}$  system and let  $\mu_Y$  be an ergodic  $S$  invariant Borel probability measure on  $Y$ . For all  $G \in L^1(Y, \mu_Y)$ , there exists a set  $Y' \subset Y$  with  $\mu_Y(Y') = 1$  such that for any ergodic  $\mathbb{Z}$  system  $(X, T, \mu_X)$  which is Kronecker disjoint from  $(Y, S, \mu_Y)$ , any  $k \in \mathbb{N}$ , any  $f_1, \dots, f_k \in L^\infty(X, \mu_X)$ , and any  $y \in Y'$ ,*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N G(S^n y) \prod_{i=1}^k T^{in} f_i = \left( \int_Y G \, d\mu_Y \right) \cdot \left( \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k T^{in} f_i \right)$$

in  $L^2(X, \mu_X)$ . Moreover, if  $(Y, S)$  is uniquely ergodic and  $G \in C(Y)$ , then we can take  $Y' = Y$ .

**1.1. Structure of the paper.** In §2, we review basic results and facts regarding Kronecker factors and distal systems, which are needed in the subsequent sections. In §3, we discuss Berg’s notion of quasi-disjointness for  $\mathbb{Z}$  systems in more detail and give a proof of Theorem 1.3. In §4, we provide a proof of Theorem 1.4 by showing that quasi-disjointness lifts through group extensions, is preserved by passing to factors and is preserved under taking inverse limits. Sections 5 and 6 contain numerous applications of our main results to questions about pointwise convergence in ergodic theory and to the theory of multiple recurrence, including proofs of Theorems 1.7 and 1.8. Finally, in §7 we formulate some natural open questions.

**2. Preliminaries**

In this section, we present various preliminary results, which will be of use throughout the paper, on  $G$  systems, their Kronecker factors, and their joinings. We conclude with a brief discussion of topologically and measurably distal systems. Throughout this paper,  $G$  denotes a countable group.

**2.1. Measure-preserving systems.** By a *topological  $G$  system* we mean a pair  $(X, T)$  where  $X$  is a compact metric space and  $T$  is a continuous left action of  $G$  on  $X$ . A  *$G$  system* is a tuple  $\mathbf{X} = (X, T, \mu_X)$  where  $(X, T)$  is a topological  $G$  system and  $\mu_X$  is a  $T$  invariant Borel probability measure on  $X$ . The *product* of two  $G$  systems  $\mathbf{X} = (X, T, \mu_X)$  and  $\mathbf{Z} = (Z, R, \mu_Z)$  is the system  $\mathbf{X} \times \mathbf{Z} = (X \times Z, T \times R, \mu_X \otimes \mu_Z)$  where  $T \times R$  is the diagonal action  $(T \times R)^g = T^g \times R^g$ .

A  $G$  system  $\mathbf{Z} = (Z, R, \mu_Z)$  is a *factor* of a  $G$  system  $\mathbf{X} = (X, T, \mu_X)$  if there is a  $G$  invariant, conull (that is full measure) subset  $X'$  of  $X$  and a measurable, measure-preserving,  $G$  equivariant map  $X' \rightarrow Z$ . Any such map, together with its conull, invariant domain, is called a *factor map*.

Given a factor  $\mathbf{Z}$  of a  $G$  system  $\mathbf{X}$ , the associated factor map induces an isometric embedding of  $L^2(\mathbf{Z}, \mu_{\mathbf{Z}})$  in  $L^2(\mathbf{X}, \mu_{\mathbf{X}})$ . Denote by  $\mathbb{E}(\cdot|\mathbf{Z})$  the orthogonal projection from  $L^2(\mathbf{X}, \mu_{\mathbf{X}})$  to this embedded copy of  $L^2(\mathbf{Z}, \mu_{\mathbf{Z}})$ . Given  $f \in L^2(\mathbf{X}, \mu_{\mathbf{X}})$ , we can think of  $\mathbb{E}(f|\mathbf{Z})$  as a function either on  $X$  or on  $Z$ .

2.2. *Disintegrations.* Given a factor map  $\pi : \mathbf{X} \rightarrow \mathbf{Y}$  of  $G$  systems one can always find an almost-surely defined, measurable family  $y \mapsto \mu_y$  of Borel probability measures on  $X$  such that

$$\int f \, d\mu_X = \iint f \, d\mu_y \, d\mu_Y(y)$$

for all  $f$  in  $L^1(X, \mu_X)$  and that  $T^g \mu_y = \mu_{Sg y}$  for all  $g \in G$  almost surely. Moreover, the family  $y \mapsto \mu_y$  is uniquely determined almost surely by these properties. If  $\mathbf{Y}$  is the factor corresponding (via [Zim76b, Corollary 2.2]) to the  $\sigma$ -algebra of  $T$  invariant sets, then the resulting disintegration  $y \mapsto \mu_y$  is a version of the ergodic decomposition of  $\mu_X$ . We refer the reader to [EW11, Ch. 5] for details on disintegrations of measures.

2.3. *The Kronecker factor.* Every  $G$  system  $\mathbf{X} = (X, T, \mu_X)$  induces a right action of  $G$  on  $L^2(X, \mu_X)$  defined by  $(T^g f)(x) = f(T^g x)$ . A function  $f \in L^2(X, \mu_X)$  is *almost periodic* if its orbit  $\{T^g f : g \in G\}$  has compact closure in the strong topology of  $L^2(X, \mu_X)$ . We write  $\text{AP}(\mathbf{X})$  for the closed subspace of  $L^2(X, \mu_X)$  spanned by almost periodic functions. A  $G$  system  $\mathbf{X}$  is *almost periodic* if  $L^2(X, \mu_X) = \text{AP}(\mathbf{X})$ . It follows from [LG61, Lemma 4.3] that  $\text{AP}(\mathbf{X})$  coincides with the subspace of  $L^2(X, \mu_X)$  spanned by finite-dimensional,  $T$  invariant subspaces of  $L^2(X, \mu_X)$ . There is a  $G$  invariant, countably generated sub- $\sigma$ -algebra  $\mathcal{A}$  of the Borel  $\sigma$ -algebra of  $X$  such that  $\text{AP}(\mathbf{X}) = L^2(X, \mathcal{A}, \mu_X)$  (cf. [FK91, Lemma 3.1]). By [Zim76b, Corollary 2.2] the  $\sigma$ -algebra  $\mathcal{A}$  corresponds to a factor  $\mathbf{KX}$  of  $\mathbf{X}$  called the *Kronecker factor*.

PROPOSITION 2.1. *For any two  $G$  systems  $\mathbf{X}$  and  $\mathbf{Z}$ , we have  $\text{AP}(\mathbf{X} \times \mathbf{Z}) = \text{AP}(\mathbf{X}) \otimes \text{AP}(\mathbf{Z})$ .*

*Proof.* Given a  $G$  system  $\mathbf{Y}$ , denote by  $\text{WM}(\mathbf{Y})$  the closure in  $L^2(Y, \mu_Y)$  of the collection of vectors  $f$  with the property that 0 belongs to the weak closure of  $\{T^g f : g \in G\}$ . By [LG61, Corollary 4.12] one can write  $L^2(Y, \mu_Y)$  as the direct sum  $\text{AP}(\mathbf{Y}) \oplus \text{WM}(\mathbf{Y})$ .

Fix now  $G$  systems  $\mathbf{X} = (X, T, \mu_X)$  and  $\mathbf{Z} = (Z, R, \mu_Z)$ . We have

$$\begin{aligned} L^2(X \times Z, \mu_X \otimes \mu_Z) &= (\text{AP}(\mathbf{X}) \otimes \text{AP}(\mathbf{Z})) \oplus (\text{WM}(\mathbf{X}) \otimes \text{AP}(\mathbf{Z})) \\ &\quad \oplus (\text{AP}(\mathbf{X}) \otimes \text{WM}(\mathbf{Z})) \oplus (\text{WM}(\mathbf{X}) \otimes \text{WM}(\mathbf{Z})) \end{aligned}$$

holds. Certainly  $\text{AP}(\mathbf{X}) \otimes \text{AP}(\mathbf{Z}) \subset \text{AP}(\mathbf{X} \times \mathbf{Z})$ . If  $f$  belongs to  $\text{WM}(\mathbf{X})$ , then  $f \otimes g$  belongs to  $\text{WM}(\mathbf{X} \times \mathbf{Z})$  for every  $g$  in  $L^2(Z, \mu_Z)$ . Similarly, if  $g$  belongs to  $\text{WM}(\mathbf{Z})$ , then  $f \otimes g$  belongs to  $\text{WM}(\mathbf{X} \times \mathbf{Z})$  for every  $f$  in  $L^2(X, \mu_X)$ . This gives

$$(\text{WM}(\mathbf{X}) \otimes \text{AP}(\mathbf{Z})) \oplus (\text{AP}(\mathbf{X}) \otimes \text{WM}(\mathbf{Z})) \oplus (\text{WM}(\mathbf{X}) \otimes \text{WM}(\mathbf{Z})) \subset \text{WM}(\mathbf{X} \times \mathbf{Z})$$

and the result follows. □

When  $\mathbf{X}$  is ergodic [Mac64, Theorem 1] allows us to model the system  $\mathbf{KX}$  as a homogeneous space  $K/H$  where  $K$  is a compact group and  $H$  is a closed subgroup with the action of  $G$  on  $K/H$  given by a homomorphism  $G \rightarrow K$  with dense image. Write  $\mathbf{KX}$  for the underlying space of any such modeling of  $\mathbf{KX}$ . For ergodic  $\mathbb{Z}$  systems the Kronecker factor can be explicitly described by the system’s discrete spectrum.

*Definition 2.2.* Let  $\mathbf{X} = (X, T, \mu_X)$  be a  $\mathbb{Z}$  system. The *discrete spectrum* of  $\mathbf{X}$ , denoted by  $\text{Eig}(\mathbf{X})$ , is defined to be the set of all eigenvalues of  $T$  when viewed as a unitary operator  $T : L^2(X, \mu_X) \rightarrow L^2(X, \mu_X)$ ,

$$\text{Eig}(\mathbf{X}) = \{\zeta \in \mathbb{C} : \text{there exists } f \in L^2(X, \mu_X) \text{ with } f \neq 0 \text{ and } Tf = \zeta f\}.$$

Given two ergodic  $G$  systems  $\mathbf{X}$  and  $\mathbf{Z}$  we can (using the Bohr compactification of  $G$ , for instance) assume that  $\mathbf{KX}$  and  $\mathbf{KZ}$  are modeled by homogeneous spaces of the same compact topological group  $K$ . That is, we may assume  $\mathbf{KX} = K/H_X$  and  $\mathbf{KZ} = K/H_Z$  for some compact topological group  $K$  and closed subgroups  $H_X, H_Z$  thereof, the action of  $G$  on both spaces determined by a homomorphism  $G \rightarrow K$  with dense image. The system  $K/(H_X, H_Z)$  is a factor of both  $\mathbf{X}$  and  $\mathbf{Z}$  and serves as a model for  $\mathbf{K}(\mathbf{X}, \mathbf{Z})$ , their *joint Kronecker factor* by which we mean the largest almost periodic factor of both  $\mathbf{X}$  and  $\mathbf{Z}$ .

**PROPOSITION 2.3.** (Cf. [Fur81, Lemma 4.18]) *If  $\mathbf{X}$  and  $\mathbf{Z}$  are ergodic  $G$  systems and  $\mathbf{X} \times \mathbf{Z}$  is not ergodic, then  $L^2(X, \mu_X)$  and  $L^2(Z, \mu_Z)$  contain isomorphic, finite-dimensional, invariant subspaces of non-constant functions.*

*Proof.* If  $\mathbf{X} \times \mathbf{Z}$  is not ergodic, then there is a non-constant function  $f$  in  $L^2(X \times Z, \mu_X \otimes \mu_Z)$  that is  $T \times R$  invariant. Therefore,  $\langle f, \phi \rangle = \langle f, (T \times R)^g \phi \rangle$  for all  $\phi \in L^2(X \times Z, \mu_X \otimes \mu_Z)$  and  $g \in G$ . In particular, if  $\phi \in \text{WM}(\mathbf{X} \times \mathbf{Z})$ , then  $\langle f, \phi \rangle = 0$ . It follows that  $f$  belongs to  $\text{AP}(\mathbf{X} \times \mathbf{Z})$  and, therefore (by Proposition 2.1), to  $\text{AP}(\mathbf{X}) \otimes \text{AP}(\mathbf{Z})$ . Now  $\text{AP}(\mathbf{X})$  and  $\text{AP}(\mathbf{Z})$  are spanned by finite-dimensional, invariant subspaces of  $L^2(X, \mu_X)$  and  $L^2(Z, \mu_Z)$ , respectively, thus

$$f = \sum_{\iota, \kappa} g_\iota \otimes h_\kappa$$

where  $\iota$  and  $\kappa$  enumerate the finite-dimensional subrepresentations of  $L^2(X, \mu_X)$  and  $L^2(Z, \mu_Z)$ , respectively, and  $g_\iota \otimes h_\kappa$  is the projection of  $f$  on the corresponding subrepresentation of  $L^2(X \times Z, \mu_X \otimes \mu_Z)$ . The fact that  $f$  is invariant implies only terms of the form  $g_\iota \otimes h_{\iota^*}$  contribute to the sum, where  $\iota^*$  denotes the contragradient of the representation  $\iota$ . As  $f$  is non-constant there must be a non-trivial, finite-dimensional representation of  $G$  that appears as a subrepresentation of both  $L^2(X, \mu_X)$  and  $L^2(Z, \mu_Z)$ . □

**2.4. Joinings.** Given  $G$  systems  $\mathbf{X} = (X, T, \mu_X)$  and  $\mathbf{Z} = (Z, R, \mu_Z)$ , a measure  $\lambda$  on  $X \times Z$  is a *joining* of  $\mathbf{X}$  and  $\mathbf{Z}$  if  $\lambda$  is invariant under the diagonal action  $T \times R$  and the pushforwards of  $\lambda$  under the two coordinate projection maps  $\pi_X : X \times Z \rightarrow X$  and  $\pi_Z : X \times Z \rightarrow Z$  satisfy  $\pi_X(\lambda) = \mu_X$  and  $\pi_Z(\lambda) = \mu_Z$ . We write  $\mathcal{J}(\mathbf{X}, \mathbf{Z})$  for the set of all joinings of  $\mathbf{X}$  with  $\mathbf{Z}$  and  $\mathcal{J}_e(\mathbf{X}, \mathbf{Z})$  for the set of joinings of  $\mathbf{X}$  with  $\mathbf{Z}$  that are ergodic. The product  $\mu_X \otimes \mu_Z$  is always a joining of  $\mathbf{X}$  and  $\mathbf{Z}$ . When  $\mathbf{X}$  and  $\mathbf{Z}$  are ergodic, the



set  $\mathcal{J}_e(\mathbf{X}, \mathbf{Z})$  is always non-empty because the measures in the ergodic decomposition of  $\mu_X \otimes \mu_Z$  can be shown to be ergodic joinings of  $\mathbf{X}$  and  $\mathbf{Z}$ .

One says that  $G$  systems  $\mathbf{X}$  and  $\mathbf{Z}$  are *disjoint* if  $\mu_X \otimes \mu_Z$  is their only joining. We say that  $G$  systems  $\mathbf{X}$  and  $\mathbf{Z}$  are *Kronecker disjoint* if their Kronecker factors  $\mathbf{KX}$  and  $\mathbf{KZ}$  are disjoint.

**2.5. Distal systems.** Given a  $G$  system  $\mathbf{X} = (X, T, \mu_X)$  denote by  $\text{Aut}(\mathbf{X})$  the group of invertible, measurable, measure-preserving maps on  $(X, \mu_X)$  that commute with  $T$ , where two automorphisms are identified if they coincide  $\mu_X$  almost everywhere. As  $X$  is a compact metric space, the group  $\text{Aut}(\mathbf{X})$  is metrizable and as such becomes a Polish topological group. Given a compact subgroup  $L$  of  $\text{Aut}(\mathbf{X})$ , the associated sub- $\sigma$ -algebra of  $L$  invariant sets determines a factor of  $\mathbf{X}$ . One says that a  $G$  system  $\mathbf{X}$  is a *group extension* of a  $G$  system  $\mathbf{Y}$  if  $\mathbf{Y}$  is (isomorphic to) a factor of  $\mathbf{X}$  via a compact subgroup of  $\text{Aut}(\mathbf{X})$  in the above fashion. As stated in the introduction, a  $G$  system is *measurably distal* if it belongs to the smallest class of  $G$  systems that contains the trivial one-point system and is closed under group extensions, factor maps, and inverse limits.

We recall that a topological  $G$  system  $(X, T)$  is *topologically distal* if

$$\inf\{d(T^g x, T^g x') : g \in G\} > 0$$

for all  $x \neq x'$  in  $X$ . In the case  $G = \mathbb{Z}$  Parry [Par68] modified this definition to apply to measure-preserving actions. We now recall Zimmer's generalization [Zim76a, Definition 8.5] of Parry's definition to actions of countable groups. Given a  $G$  system  $\mathbf{X} = (X, T, \mu_X)$  a sequence  $n \mapsto A_n$  of Borel subsets of  $X$  with  $\mu_X(A_n) > 0$  and  $\mu_X(A_n) \rightarrow 0$  is called a *separating sieve* if there is a conull set  $X' \subset X$  such that, whenever  $x, x' \in X'$  and, for each  $n \in \mathbb{N}$ , one can find  $g_n \in G$  with  $\{T^{g_n} x, T^{g_n} x'\} \subset A_n$ , one has  $x = x'$ . Zimmer [Zim76a, Theorem 8.7] proved that a non-atomic  $G$  system  $\mathbf{X}$  is measurably distal if and only if it has a separating sieve. Using this characterization of distality, one can show that every topological  $G$  system  $(X, T)$  that is topologically distal has the property that for every  $T$  invariant Borel probability measure  $\mu_X$  on  $X$  the  $G$  system  $(X, T, \mu_X)$  is measurably distal. Lindenstrauss [Lin99] has proved a partial converse to this result for  $\mathbb{Z}$  systems by showing that every measurably distal  $\mathbb{Z}$  system can be modeled by a topologically distal  $\mathbb{Z}$  system equipped with an invariant Borel probability measure.

### 3. Proof of Theorem 1.3

In this section, we prove Theorem 1.3. It follows a preparatory discussion parameterizing the space of joinings of ergodic, almost-periodic  $\mathbb{Z}$  systems and describing the ergodic decomposition of the product of two such  $\mathbb{Z}$  systems.

As described in §2.3, the Kronecker factor of an ergodic  $\mathbb{Z}$  system can be modeled as an ergodic rotation on a compact abelian group. Given two ergodic rotations on compact, abelian groups  $\mathbf{X} = (X, T, \mu_X)$  and  $\mathbf{Y} = (Y, S, \mu_Y)$ , their ergodic joinings can be easily described as follows: let  $e_X$  be the identity of the compact abelian group  $X$ , let  $e_Y$  be the identity of  $Y$ , and let  $H$  be the subgroup

$$H = \overline{\{(T \times S)^n(e_X, e_Y) : n \in \mathbb{Z}\}}$$



of  $X \times Y$ . Given an ergodic joining  $\lambda \in \mathcal{J}_e(\mathbf{X}, \mathbf{Y})$ , let  $(x_0, y_0) \in X \times Y$  be a generic point. The support of  $\lambda$  is the orbit closure of  $(x_0, y_0)$ , which is  $(x_0, y_0) + H$ . Hence, the pushforward of  $\lambda$  under the map  $(x, y) \mapsto (x - x_0, y - y_0)$  is a measure on  $H$  invariant under  $T \times S$  and, hence, must be the Haar measure on  $H$ . It follows that  $\lambda$  is the Haar measure on the coset  $(x_0, y_0) + H$ .

Now let  $K = (X \times Y)/H$  be the group of cosets of  $H$  and write  $\mu_K$  for Haar measure on  $K$ . Define a map  $R : K \rightarrow K$  by  $R : (x, y) + H \mapsto (Tx, y) + H = (x, S^{-1}y) + H$ . Let  $\alpha : X \rightarrow K$  be defined by  $\alpha(x) = (x, e_Y) + H$  and let  $\beta : Y \rightarrow K$  be defined by  $\beta(y) = (e_X, -y) + H$ . It is easy to check that both  $\alpha$  and  $\beta$  are factor maps onto  $(K, R)$  and, hence, either  $\mathbf{X} \times \mathbf{Y}$  is ergodic (and, thus,  $K = \{\text{id}\}$ ), or  $\mathbf{X}$  and  $\mathbf{Y}$  share a non-trivial common factor.

Note that  $\gamma(x, y) = (x, y) + H$  from  $\mathbf{X} \times \mathbf{Y}$  to  $K$  is the maximal invariant factor (because every invariant function is constant along cosets of  $H$ ). In fact,  $(K, R)$  is the maximal common factor of  $\mathbf{X}$  and  $\mathbf{Y}$ .

As  $\gamma^{-1}((x, y) + H) = (x, y) + H$ , there exists exactly one ergodic joining living in that pre-image, namely, the Haar measure. For each  $k \in K$ , let  $\lambda_k \in \mathcal{J}_e(\mathbf{X}, \mathbf{Y})$  be the unique joining such that  $\lambda_k(\gamma^{-1}(k)) = 1$ . Observe that

$$\mu_X \otimes \mu_Y = \int_K \lambda_k d\mu_K(k) \tag{2}$$

is, therefore, the ergodic decomposition of  $\mu_X \otimes \mu_Y$ .

*Proof of Theorem 1.3.* First, suppose that  $\mathbf{X}$  and  $\mathbf{Y}$  are quasi-disjoint. Let  $(K, \mu_K)$  be a model for their joint Kronecker factor and let  $k \mapsto \lambda_k$  be a measurable map from  $K$  into the space  $\mathcal{J}_e(\mathbf{X}, \mathbf{Y})$  of ergodic joinings of  $\mathbf{X}$  and  $\mathbf{Y}$  such that  $\lambda_k$  gives full measure to  $\gamma^{-1}(k)$  for almost every  $k$ . Then

$$\eta = \int \lambda_k d\mu_K(k)$$

is a joining of  $\mathbf{X}$  and  $\mathbf{Y}$ . We claim that the projection of  $\eta$  to  $K\mathbf{X} \times K\mathbf{Y}$  is the product measure. Indeed, let  $\pi : \mathbf{X} \times \mathbf{Y} \rightarrow K\mathbf{X} \times K\mathbf{Y}$  be the corresponding factor map. We can decompose  $\gamma = \tilde{\gamma} \circ \pi$  for some  $\tilde{\gamma} : K\mathbf{X} \times K\mathbf{Y} \rightarrow K$ . As  $\pi(\lambda_k)$  gives full measure to  $\tilde{\gamma}^{-1}(k)$ , the discussion preceding this proof implies it is uniquely determined by  $k$ . In particular, in view of equation 2, we have

$$\pi(\eta) = \int \pi(\lambda_k) d\mu_K(k) = \mu_{K\mathbf{X}} \otimes \mu_{K\mathbf{Y}}$$

as claimed. However, then by quasi-disjointness  $\eta = \mu_{\mathbf{X}} \otimes \mu_{\mathbf{Y}}$ . By uniqueness of the ergodic disintegration, the map  $k \rightarrow \lambda_k$  is uniquely defined almost everywhere, so  $\mathbf{X}$  and  $\mathbf{Y}$  satisfy (BQD).

Conversely, suppose that  $\mathbf{X}$  and  $\mathbf{Y}$  satisfy (BQD). Let  $\eta$  be a joining of  $\mathbf{X}$  and  $\mathbf{Y}$  that projects to the product measure on  $K\mathbf{X} \times K\mathbf{Y}$ . Then  $\gamma\eta$  is the Haar measure on the maximal common Kronecker factor  $K(\mathbf{X}, \mathbf{Y})$  of  $\mathbf{X}$  and  $\mathbf{Y}$ . The ergodic disintegration of  $\eta$  is the same as the disintegration of  $\eta$  over  $K(\mathbf{X}, \mathbf{Y})$ . By (BQD), this disintegration is, in turn, the same as the disintegration of  $\mu \otimes \nu$  over  $K(\mathbf{X}, \mathbf{Y})$ . Therefore,  $\eta = \mu \otimes \nu$ . □

4. *Quasi-disjointness for measurably distal systems*

In this section, we give a proof of Theorem 1.4. The proof comprises three parts, covered in the following three subsections. The first part consists of showing that quasi-disjointness lifts through group extensions. The second part proves that quasi-disjointness is preserved when passing to a factor and the third part consists of showing that quasi-disjointness is preserved by inverse limits. As, starting from the trivial system, such operations exhaust the class of distal systems, these three parts combined indeed yield a complete proof of Theorem 1.4. We conclude this section with an example of a  $\mathbb{Z}$  system that is not measurably distal but is quasi-disjoint from every ergodic system.

4.1. *Quasi-disjointness lifts through group extensions.* The purpose of this subsection is to prove the following result.

**THEOREM 4.1.** *Let  $\mathbf{X} = (X, T, \mu_X)$ ,  $\mathbf{Y} = (Y, S, \mu_Y)$ , and  $\mathbf{Z} = (Z, R, \mu_Z)$  be  $G$  systems and assume that  $\mathbf{X}$  is a group extension of  $\mathbf{Y}$ . If  $\mathbf{Y}$  is quasi-disjoint from  $\mathbf{Z}$ , then so is  $\mathbf{X}$ .*

For the proof of Theorem 4.1, we borrow ideas from the proof of [Fur67, Theorem 1.4]. See [Gla03, Theorem 3.30] for a version of that result applicable to actions of countable groups.

*Proof of Theorem 4.1.* Let  $\pi_{\mathbf{X},\mathbf{KX}}$  and  $\pi_{\mathbf{Z},\mathbf{KZ}}$  denote the projection maps from  $\mathbf{X}$  onto  $\mathbf{KX}$  and from  $\mathbf{Z}$  onto  $\mathbf{KZ}$ , respectively. Let  $\lambda \in \mathcal{J}(\mathbf{X}, \mathbf{Z})$  be a joining of  $\mathbf{X}$  and  $\mathbf{Z}$  with the property that  $(\pi_{\mathbf{X},\mathbf{KX}} \times \pi_{\mathbf{Z},\mathbf{KZ}})(\lambda) = \mu_{\mathbf{KX}} \otimes \mu_{\mathbf{KZ}}$ . We want to show that  $\lambda = \mu_X \otimes \mu_Z$ .

Let  $\pi_{\mathbf{X},\mathbf{Y}}$  denote the factor map from  $\mathbf{X}$  onto  $\mathbf{Y}$ . Note that  $(\pi_{\mathbf{X},\mathbf{Y}} \times \text{id}_{\mathbf{Z}})(\lambda)$  is a joining of  $\mathbf{Y}$  with  $\mathbf{Z}$  whose projection onto the product of the Kronecker factors  $\mathbf{KY} \times \mathbf{KZ}$  equals  $\mu_{\mathbf{KY}} \otimes \mu_{\mathbf{KZ}}$ . As  $\mathbf{Y}$  is quasi-disjoint from  $\mathbf{Z}$ , we conclude that  $(\pi_{\mathbf{X},\mathbf{Y}} \times \text{id}_{\mathbf{Z}})(\lambda) = \mu_Y \otimes \mu_Z$ .

As  $\mathbf{X}$  is a group extension of  $\mathbf{Y}$ , there exists a compact group  $L \leq \text{Aut}(\mathbf{X})$  such that the factor  $\mathbf{Y}$  corresponds to the sub- $\sigma$ -algebra of  $L$  invariant subsets of  $\mathbf{X}$  (cf. §2.5). Let  $\mu_L$  denote the normalized Haar measure on  $L$  and let  $\mathbf{L} = (L, Q, \mu_L)$ , where  $Q$  is the action of  $L$  on itself by left multiplication. Given  $\psi \in L^\infty(L, \mu_L)$  with  $\psi \geq 0$  we define a new joining  $\lambda_\psi \in \mathcal{J}(\mathbf{X}, \mathbf{Z})$  by

$$\lambda_\psi(A) = \int_{X \times Z} \int_K 1_A(lx, z) \psi(l) d\mu_L(l) d\lambda(x, z) \tag{3}$$

for all Borel sets  $A \subset X \times Z$ . Let  $\lambda_1$  denote the measure defined by (3) with  $\psi = 1$ . We claim that  $\lambda_1 = \mu_X \otimes \mu_Z$ . To verify this claim, let  $f \in L^2(X, \mu_X)$ ,  $g \in L^2(Z, \mu_Z)$ , and define  $f'(x) = \int_L f(lx) d\mu_L(l)$  for all  $x \in X$ . Note that  $f'$  is  $L$  invariant and, hence, there exists  $h \in L^2(Y, \mu_Y)$  such that  $h \circ \pi_{\mathbf{X},\mathbf{Y}} = f'$ . We have

$$\begin{aligned} \int_{X \times Z} f \otimes g d\lambda_1 &= \int_{X \times Z} f' \otimes g d\lambda \\ &= \int_{X \times Z} (h \otimes g) \circ (\pi_{\mathbf{X},\mathbf{Y}} \times \text{id}_{\mathbf{Z}}) d\lambda \\ &= \int_{Y \times Z} h \otimes g d(\pi_{\mathbf{X},\mathbf{Y}} \times \text{id}_{\mathbf{Z}})(\lambda) \end{aligned}$$

$$\begin{aligned}
 &= \int_{Y \times Z} h \otimes g \, d\mu_Y \otimes \mu_Z \\
 &= \int_{X \times Z} f' \otimes g \, d\mu_X \otimes \mu_Z = \int_{X \times Z} f \otimes g \, d\mu_X \otimes \mu_Z
 \end{aligned}$$

because  $(\pi_{\mathbf{X},\mathbf{Y}} \times \text{id}_{\mathbf{Z}})\lambda = \mu_Y \otimes \mu_Z$ . This shows that indeed  $\lambda_1 = \mu_X \otimes \mu_Z$ .

Next, observe that

$$\lambda_\psi(A) \leq \lambda_1(A)\|\psi\|_\infty = (\mu_X \otimes \mu_Z)(A)\|\psi\|_\infty. \tag{4}$$

for all Borel sets  $A \subset X \times Z$ . Inequality (4) shows that  $\lambda_\psi$  is absolutely continuous with respect to  $\mu_X \otimes \mu_Z$ . Let  $F_\psi$  denote the Radon–Nikodym derivative of  $\lambda_\psi$  with respect to  $\mu_X \otimes \mu_Z$ . It also follows from (4) that  $\|F_\psi\|_\infty \leq \|\psi\|_\infty$  and so  $F_\psi \in L^\infty(X \times Z, \mu_X \otimes \mu_Z)$ . Moreover, since  $\lambda_\psi$  is a  $T \times R$  invariant measure, we conclude that  $F_\psi$  is a  $T \times R$  invariant function in  $L^\infty(X \times Z, \mu_X \otimes \mu_Z)$ . As any  $T \times R$  invariant function is almost periodic, it follows that  $F_\psi \in \text{AP}(\mathbf{X} \times \mathbf{Z})$ . Therefore,  $F_\psi \in \text{AP}(\mathbf{X}) \otimes \text{AP}(\mathbf{Z})$  by Proposition 2.1. It follows that for all  $f \in L^2(X, \mu_X)$  and  $g \in L^2(Z, \mu_Z)$ ,

$$\begin{aligned}
 \int_{X \times Z} f \otimes g \, d\lambda_\psi &= \int_{X \times Z} F_\psi \cdot (f \otimes g) \, d(\mu_X \otimes \mu_Z) \\
 &= \int_{X \times Z} F_\psi \cdot (\mathbb{E}(f|\mathbf{K}\mathbf{X}) \otimes \mathbb{E}(g|\mathbf{K}\mathbf{Z})) \, d(\mu_X \otimes \mu_Z) \\
 &= \int_{X \times Z} \mathbb{E}(f|\mathbf{K}\mathbf{X}) \otimes \mathbb{E}(g|\mathbf{K}\mathbf{Z}) \, d\lambda_\psi \\
 &= \int_{X \times Z} \int_L \mathbb{E}(f|\mathbf{K}\mathbf{X})(lx) \mathbb{E}(g|\mathbf{K}\mathbf{Z})(z) \psi(l) \, d\mu_L(l) \, d\lambda(x, z) \\
 &= \int_L \left( \int_{X \times Z} \mathbb{E}(lf|\mathbf{K}\mathbf{X})(x) \mathbb{E}(g|\mathbf{K}\mathbf{Z})(z) \, d\lambda(x, z) \right) \psi(l) \, d\mu_L(l).
 \end{aligned}$$

As  $(\pi_{\mathbf{X},\mathbf{K}\mathbf{X}} \times \pi_{\mathbf{Z},\mathbf{K}\mathbf{Z}})(\lambda) = \mu_{\mathbf{K}\mathbf{X}} \otimes \mu_{\mathbf{K}\mathbf{Z}}$  we have

$$\begin{aligned}
 \int_{X \times Z} \mathbb{E}(lf|\mathbf{K}\mathbf{X}) \otimes \mathbb{E}(g|\mathbf{K}\mathbf{Z}) \, d\lambda &= \int_{X \times Z} \mathbb{E}(lf|\mathbf{K}\mathbf{X}) \otimes \mathbb{E}(g|\mathbf{K}\mathbf{Z}) \, d(\mu_X \otimes \mu_Z) \\
 &= \int_{X \times Z} f \otimes g \, d(\mu_X \otimes \mu_Z).
 \end{aligned}$$

for all  $l \in L$ . We conclude that

$$\int_{X \times Z} f \otimes g \, d\lambda_\psi = \left( \int_{X \times Z} f \otimes g \, d(\mu_X \otimes \mu_Z) \right) \left( \int_L \psi(l) \, d\mu_L(l) \right)$$

so, in particular, for any  $\psi \in L^\infty(L, \mu_L)$  with  $\psi \geq 0$  and  $\int_L \psi \, d\mu_L = 1$ , the measure  $\lambda_\psi$  coincides with  $\mu_X \otimes \mu_Z$ .

Finally, allowing  $\psi$  to run through an approximate identity, one can approximate  $\lambda$  by  $\lambda_\psi$  and thereby conclude that  $\lambda = \mu_X \otimes \mu_Z$ , which finishes the proof.  $\square$

4.2. *Quasi-disjointness passes to factors.* In this subsection, we prove the following theorem.

**THEOREM 4.2.** *Let  $\mathbf{X} = (X, T, \mu_X)$ ,  $\mathbf{Y} = (Y, S, \mu_Y)$ , and  $\mathbf{Z} = (Z, R, \mu_Z)$  be  $G$  systems and suppose that  $\mathbf{Y}$  is a factor of  $\mathbf{X}$ . If  $\mathbf{X}$  and  $\mathbf{Z}$  are quasi-disjoint, then  $\mathbf{Y}$  and  $\mathbf{Z}$  are also quasi-disjoint.*

For the proof of Theorem 4.2, we need to recall the definition of relatively independent joinings. Let  $\mathbf{X} = (X, T, \mu_X)$ ,  $\mathbf{Y} = (Y, S, \mu_Y)$ , and  $\mathbf{Z} = (Z, R, \mu_Z)$  be  $G$  systems and suppose that  $\mathbf{Y}$  is a factor of both  $\mathbf{X}$  and  $\mathbf{Z}$ . Let  $\pi_{\mathbf{X},\mathbf{Y}} : X \rightarrow Y$  and  $\pi_{\mathbf{Z},\mathbf{Y}} : Z \rightarrow Y$  denote the respective factor maps. Using  $\pi_{\mathbf{X},\mathbf{Y}}$ , we can embed  $L^2(Y, \mu_Y)$  into  $L^2(X, \mu_X)$ . Likewise, through  $\pi_{\mathbf{Z},\mathbf{Y}}$  we can identify  $L^2(Y, \mu_Y)$  as a subspace of  $L^2(Z, \mu_Z)$ . The relatively independent joining of  $\mathbf{X}$  with  $\mathbf{Z}$  over  $\mathbf{Y}$  is the triple  $\mathbf{X} \times_{\mathbf{Y}} \mathbf{Z} = (X \times Z, T \times R, \mu_X \otimes_{\mathbf{Y}} \mu_Z)$ , where  $\mu_X \otimes_{\mathbf{Y}} \mu_Z$  denotes the unique measure on  $X \times Z$  with the property that

$$\int_{X \times Z} f \otimes g \, d(\mu_X \otimes_{\mathbf{Y}} \mu_Z) = \int_Y \mathbb{E}(f|\mathbf{Y})\mathbb{E}(g|\mathbf{Y}) \, d\mu_Y \tag{5}$$

for all  $f \in L^2(X, \mu_X)$  and all  $g \in L^2(Z, \mu_Z)$ .

**LEMMA 4.3.** *Let  $\mathbf{X} = (X, T, \mu_X)$  and  $\mathbf{Y} = (Y, S, \mu_Y)$  be  $G$  systems and suppose that  $\mathbf{Y}$  is a factor of  $\mathbf{X}$ . Then the relatively independent joining  $\mathbf{KX} \times_{\mathbf{KY}} \mathbf{Y}$  is a factor of  $\mathbf{X}$ .*

*Proof.* Let  $\pi_{\mathbf{X},\mathbf{Y}} : X \rightarrow Y$  denote the factor map from  $\mathbf{X}$  to  $\mathbf{Y}$  and let  $\pi_{\mathbf{X},\mathbf{KX}} : X \rightarrow \mathbf{KX}$  denote the factor map from  $\mathbf{X}$  to  $\mathbf{KX}$ . Let  $\tau : X \rightarrow \mathbf{KX} \times Y$  be defined as  $\tau(x) = (\pi_{\mathbf{X},\mathbf{KX}}(x), \pi_{\mathbf{X},\mathbf{Y}}(x))$  for all  $x \in X$ . We claim that  $\tau$  is a factor map from  $\mathbf{X}$  onto  $\mathbf{KX} \times_{\mathbf{KY}} \mathbf{Y}$ . Once this claim is verified, the proof is complete.

To show that  $\tau$  is a factor map from  $\mathbf{X}$  onto  $\mathbf{KX} \times_{\mathbf{KY}} \mathbf{Y}$ , we must show that the pushforward of  $\mu_X$  under  $\tau$  equals  $\mu_{\mathbf{KX}} \otimes_{\mathbf{KY}} \mu_Y$ . It suffices to show that

$$\int_{\mathbf{KX} \times Y} f \otimes g \, d(\tau\mu_X) = \int_{\mathbf{KX} \times Y} f \otimes g \, d\mu_{\mathbf{KX}} \otimes_{\mathbf{KY}} \mu_Y \tag{6}$$

for all  $f \in L^2(\mathbf{KX}, \mu_{\mathbf{K}})$  and all  $g \in L^2(Y, \mu_Y)$ .

By definition, the right-hand side of (6) equals  $\int_{\mathbf{KY}} \mathbb{E}(f|\mathbf{KY})\mathbb{E}(g|\mathbf{KY}) \, d\mu_{\mathbf{KY}}$ , which can be rewritten as

$$\int_X \mathbb{E}(f \circ \pi_{\mathbf{X},\mathbf{KX}}|\mathbf{KY})\mathbb{E}(g \circ \pi_{\mathbf{X},\mathbf{Y}}|\mathbf{KY}) \, d\mu_X.$$

The left-hand side of (6) equals

$$\int_X (f \circ \pi_{\mathbf{X},\mathbf{KX}})(g \circ \pi_{\mathbf{X},\mathbf{Y}}) \, d\mu_X.$$

As  $f \circ \pi_{\mathbf{X},\mathbf{KX}} = \mathbb{E}(f \circ \pi_{\mathbf{X},\mathbf{KX}}|\mathbf{KX})$  and  $g \circ \pi_{\mathbf{X},\mathbf{Y}} = \mathbb{E}(g \circ \pi_{\mathbf{X},\mathbf{Y}}|\mathbf{Y})$  we have that

$$\int_X (f \circ \pi_{\mathbf{X},\mathbf{KX}})(g \circ \pi_{\mathbf{X},\mathbf{Y}}) \, d\mu_X = \int_X \mathbb{E}(f \circ \pi_{\mathbf{X},\mathbf{KX}}|\mathbf{KX})\mathbb{E}(g \circ \pi_{\mathbf{X},\mathbf{Y}}|\mathbf{Y}) \, d\mu_X.$$

Hence, (6) is equivalent to

$$\int_X \mathbb{E}(f \circ \pi_{\mathbf{X},\mathbf{KX}}|\mathbf{KX})\mathbb{E}(g \circ \pi_{\mathbf{X},\mathbf{Y}}|\mathbf{Y}) \, d\mu_X = \int_X \mathbb{E}(f \circ \pi_{\mathbf{X},\mathbf{KX}}|\mathbf{KY})\mathbb{E}(g \circ \pi_{\mathbf{X},\mathbf{Y}}|\mathbf{KY}) \, d\mu_X. \tag{7}$$

However, because  $\mathbb{E}(\cdot|K\mathbf{X})$  and  $\mathbb{E}(\cdot|Y)$  are orthogonal projections onto  $AP(\mathbf{X})$  and  $L^2(Y, \mu_Y)$ , respectively, it follows immediately from  $AP(\mathbf{X}) \cap L^2(Y, \mu_Y) = AP(\mathbf{Y})$  that (7) is true. □

LEMMA 4.4. *Let  $\mathbf{A} = (A, T_A, \mu_A)$ ,  $\mathbf{B} = (B, T_B, \mu_B)$ ,  $\mathbf{C} = (C, T_C, \mu_C)$ , and  $\mathbf{D} = (D, T_D, \mu_D)$  be  $G$  systems. If  $\mathbf{B}$  is a factor of  $\mathbf{A}$  and  $\mathbf{D}$  is a factor of  $\mathbf{C}$ , then the induced map from  $\mathcal{J}(\mathbf{A}, \mathbf{C})$  to  $\mathcal{J}(\mathbf{B}, \mathbf{D})$  is surjective.*

*Proof.* Write  $\pi_{\mathbf{A},\mathbf{B}}$  and  $\pi_{\mathbf{C},\mathbf{D}}$  for the factor maps. Let  $b \mapsto \mu_{A,b}$  and  $d \mapsto \mu_{C,d}$  be disintegrations (cf. §2.2) of  $\mu_A$  and  $\mu_C$  over  $\mathbf{B}$  and  $\mathbf{D}$ , respectively. We have  $T_A^g \mu_{A,b} = \mu_{A,T_B^g b}$  and  $T_C^g \mu_{C,d} = \mu_{C,T_D^g d}$  almost surely for all  $g \in G$ .

Fix a joining  $\lambda_{B,D}$  of  $\mathbf{B}$  and  $\mathbf{D}$ . We claim that

$$\lambda_{A,C} = \int \mu_{A,b} \otimes \mu_{C,d} d\lambda_{B,D}(b, d)$$

is a joining of  $\mathbf{A}$  and  $\mathbf{C}$  with  $(\pi_{\mathbf{A},\mathbf{B}} \times \pi_{\mathbf{C},\mathbf{D}})\lambda_{A,C} = \lambda_{B,D}$ . First note that

$$\lambda_{A,C}(E \times C) = \int \mu_{A,b}(E) d\lambda_{B,D}(b, d) = \int \mu_{A,b}(E) d\mu_B(b) = \mu_A(E)$$

for all Borel sets  $E \subset A$  because  $\lambda_{B,D}$  is a joining, so the left marginal of  $\lambda_{A,C}$  is  $\mu_A$ . Similarly, its right marginal is  $\mu_C$ .

For all  $f \in C(B)$  and all  $h \in C(D)$  we have

$$\begin{aligned} \iint f \otimes h d(\mu_{A,b} \otimes \mu_{C,d}) d\lambda_{B,D}(b, d) &= \int \int f d\mu_{A,b} \int h d\mu_{C,d} d\lambda_{B,D}(b, d) \\ &= \int f \otimes h d\lambda_{B,D} \end{aligned}$$

by disintegration properties, so  $(\pi_{\mathbf{A},\mathbf{B}} \times \pi_{\mathbf{C},\mathbf{D}})\lambda_{A,C} = \lambda_{B,D}$ .

Finally, for any  $f \in C(A)$  and any  $h \in C(C)$ , we calculate that

$$\begin{aligned} &\iint T_A^g f \otimes T_C^g h d(\mu_{A,b} \otimes \mu_{C,d}) d\lambda_{B,D}(b, d) \\ &= \iint f \otimes h d(\mu_{A,T_B^g b} \otimes \mu_{C,T_D^g d}) d\lambda_{B,D}(b, d) \\ &= \iint f \otimes h d(\mu_{A,b} \otimes \mu_{C,d}) d\lambda_{B,D}(b, d) \\ &= \int f \otimes h d\lambda_{A,C} \end{aligned}$$

so  $\lambda_{A,C}$  is  $T_A \times T_C$  invariant. □

From Lemma 4.4, we obtain the following immediate corollary.

**COROLLARY 4.5.** *Let  $\mathbf{A} = (A, T_A, \mu_A)$ ,  $\mathbf{B} = (B, T_B, \mu_B)$ , and  $\mathbf{C} = (C, T_C, \mu_C)$  be  $G$  systems and suppose that  $\mathbf{B}$  is a factor of  $\mathbf{A}$ . Let  $\pi_{\mathbf{A},\mathbf{B}}$  denote the factor map from  $\mathbf{A}$  onto  $\mathbf{B}$ . Then, for any joining  $\lambda \in \mathcal{J}(\mathbf{B}, \mathbf{C})$ , there exists a joining  $\lambda' \in \mathcal{J}(\mathbf{A}, \mathbf{C})$  such that the pushforward of  $\lambda'$  under the factor map  $\pi_{\mathbf{A},\mathbf{B}} \times \text{id}_{\mathbf{C}}$  equals  $\lambda$ .*

$$\begin{array}{ccc} \mathbf{A} & & \mathbf{A} \times \mathbf{C} \\ \pi_{\mathbf{A},\mathbf{B}} \downarrow & & \downarrow \pi_{\mathbf{A},\mathbf{B}} \times \text{id}_{\mathbf{C}} \\ \mathbf{B} & & \mathbf{B} \times \mathbf{C} \end{array}$$

We also need the following lemmas for the proof of Theorem 4.2.

**LEMMA 4.6.** *Let  $\pi : X \rightarrow Y$  be a factor map of  $G$  systems. If  $f \in \text{AP}(Y)$ , then  $f \circ \pi \in \text{AP}(X)$ .*

*Proof.* This follows from the fact that  $\pi : L^2(Y, \mu_Y) \rightarrow L^2(X, \mu_X)$  is an isometric embedding. □

**LEMMA 4.7.** *Let  $\pi : X \rightarrow Y$  be a factor map of  $G$  systems. Let  $f \in \text{WM}(Y)$  and define  $h = f \circ \pi$ . Then  $h \in \text{WM}(X)$ .*

*Proof.* We need to prove that 0 belongs to the weak closure of  $\{T^g h : g \in G\}$  in  $L^2(X, \mu_X)$ . Note that  $\mathbb{E}(T^g h | Y) = T^g \mathbb{E}(h | Y) = T^g h$ . Fix  $\xi$  in  $L^2(X, \mu_X)$ . We have

$$\langle T^g h, \xi \rangle = \langle \mathbb{E}(T^g h | Y), \mathbb{E}(\xi | Y) \rangle,$$

so the fact that  $f \in \text{WM}(Y)$  implies that  $\{T^g h : g \in G\}$  contains 0 in its closure. □

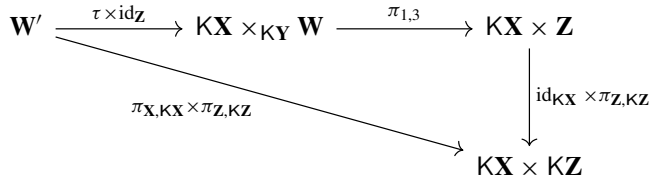
*Proof of Theorem 4.2.* Let  $\pi_{\mathbf{Y},\mathbf{KY}}$  and  $\pi_{\mathbf{Z},\mathbf{KZ}}$  denote the factor maps from  $\mathbf{Y}$  onto  $\mathbf{KY}$  and from  $\mathbf{Z}$  onto  $\mathbf{KZ}$ , respectively. Let  $\lambda \in \mathcal{J}(\mathbf{Y}, \mathbf{Z})$  be a joining of  $\mathbf{Y}$  and  $\mathbf{Z}$  with the property that  $(\pi_{\mathbf{Y},\mathbf{KY}} \times \pi_{\mathbf{Z},\mathbf{KZ}})(\lambda) = \mu_{\mathbf{KY}} \otimes \mu_{\mathbf{KZ}}$ . Let  $\mathbf{W} = (Y \times Z, S \times R, \lambda)$ . Our goal is to show that  $\mathbf{W} = \mathbf{Y} \times \mathbf{Z}$  or, equivalently, that  $\lambda = \mu_Y \otimes \mu_Z$ .

Observe that  $\mathbf{KY}$  is a factor of both  $\mathbf{W}$  and  $\mathbf{KX}$ . Hence, we can consider the relatively independent joining  $\mathbf{KX} \times_{\mathbf{KY}} \mathbf{W}$  with corresponding measure  $\mu_{\mathbf{KX}} \otimes_{\mathbf{KY}} \lambda$ . Note that the underlying space of  $\mathbf{KX} \times_{\mathbf{KY}} \mathbf{W}$  is  $\mathbf{KX} \times Y \times Z$ . Let  $\pi_{1,2} : \mathbf{KX} \times Y \times Z \rightarrow \mathbf{KX} \times Y$  denote the projection onto the first and second coordinates and let  $\pi_3 : \mathbf{KX} \times Y \times Z \rightarrow Z$  denote the projection onto the third coordinate. Observe that  $\pi_3$  is a factor map from  $\mathbf{KX} \times_{\mathbf{KY}} \mathbf{W}$  onto  $\mathbf{Z}$  and  $\pi_{1,2}$  is a factor map from  $\mathbf{KX} \times_{\mathbf{KY}} \mathbf{W}$  onto  $\mathbf{KX} \times_{\mathbf{KY}} \mathbf{Y}$ , the relatively independent joining of  $\mathbf{KX}$  with  $\mathbf{Y}$  over  $\mathbf{KY}$ . This shows that  $\mu_{\mathbf{KX}} \otimes_{\mathbf{KY}} \lambda \in \mathcal{J}(\mathbf{KX} \times_{\mathbf{KY}} \mathbf{Y}, \mathbf{Z})$ .

Let  $\tau$  denote the factor map from  $\mathbf{X}$  to  $\mathbf{KX} \times_{\mathbf{KY}} \mathbf{Y}$  in the proof of Lemma 4.3. We can now apply Corollary 4.5 with  $\mathbf{A} = \mathbf{X}$ ,  $\mathbf{B} = \mathbf{KX} \times_{\mathbf{KY}} \mathbf{Y}$ , and  $\mathbf{C} = \mathbf{Z}$  to find a joining  $\rho \in \mathcal{J}(\mathbf{X}, \mathbf{Z})$  with the property that  $(\tau \times \text{id}_{\mathbf{Z}})(\rho) = \mu_{\mathbf{KX}} \otimes_{\mathbf{KY}} \lambda$ . Let  $\mathbf{W}' = (X \times Z, T \times R, \rho)$ .

Let  $\pi_{1,3} : \mathbf{KX} \times Y \times Z \rightarrow \mathbf{KX} \times Z$  denote the projection onto the first and third coordinates. We claim that  $\pi_{1,3}(\mu_{\mathbf{KX}} \otimes_{\mathbf{KY}} \lambda) = \mu_{\mathbf{KX}} \otimes \mu_Z$ . This claim implies that the

diagram



of factor maps commutes, giving  $(\pi_{\mathbf{X},\mathbf{KX}} \times \pi_{\mathbf{Z},\mathbf{KZ}})\rho = \mu_{\mathbf{KX}} \otimes \mu_{\mathbf{KZ}}$  whence  $\rho = \mu_{\mathbf{X}} \otimes \mu_{\mathbf{Z}}$  because  $\mathbf{X}$  and  $\mathbf{Z}$  are assumed to be quasi-disjointness. It follows that the measure  $(\tau \times \text{id}_{\mathbf{Z}})(\rho)$  on  $\mathbf{KX} \times_{\mathbf{KY}} \mathbf{W}$  is the product of the measures from  $\mathbf{KX} \times_{\mathbf{KY}} \mathbf{Y}$  and  $\mathbf{Z}$  and, hence,  $\lambda = \mu_{\mathbf{Y}} \otimes \mu_{\mathbf{Z}}$  as desired.

It remains to prove the claim. Fix  $f$  in  $\text{AP}(\mathbf{X})$  and  $\phi$  in  $L^2(\mathbf{Z}, \mu_{\mathbf{Z}})$ . We have

$$\begin{aligned}
 \int f \otimes \phi \, d\pi_{1,3}(\mu_{\mathbf{KX}} \otimes_{\mathbf{KY}} \lambda) &= \int f \otimes 1 \otimes \phi \, d(\mu_{\mathbf{KX}} \otimes_{\mathbf{KY}} \lambda) \\
 &= \int \mathbb{E}(f|\mathbf{KY}) \mathbb{E}(1 \otimes \phi|\mathbf{KY}) \, d\mu_{\mathbf{KY}},
 \end{aligned}$$

so it suffices to prove that

$$\mathbb{E}(1 \otimes \phi|\mathbf{KY}) = \int \phi \, d\mu_{\mathbf{Z}} \tag{8}$$

in  $L^2(\mathbf{Y} \times \mathbf{Z}, \lambda)$ . Fix  $\psi$  in  $\text{AP}(\mathbf{Y})$ . We have

$$\int (\psi \otimes 1)(1 \otimes \phi) \, d\lambda = \int (\psi \otimes 1) \mathbb{E}(1 \otimes \phi|\mathbf{KW}) \, d\lambda$$

because  $\psi \otimes 1$  is in  $\text{AP}(\mathbf{W})$  by Lemma 4.6. However,

$$\mathbb{E}(1 \otimes \phi|\mathbf{KW}) = \mathbb{E}(1 \otimes \mathbb{E}(\phi|\mathbf{KZ})|\mathbf{KW})$$

by Lemmas 4.6 and 4.7 upon writing  $\phi = \mathbb{E}(\phi|\mathbf{KZ}) + (\phi - \mathbb{E}(\phi|\mathbf{KZ}))$ . Thus, we calculate that

$$\begin{aligned}
 \int (\psi \otimes 1)(1 \otimes \phi) \, d\lambda &= \int (\psi \otimes 1)(1 \otimes \mathbb{E}(\phi|\mathbf{KZ})) \, d\lambda \\
 &= \int \psi \otimes \mathbb{E}(\phi|\mathbf{KZ}) \, d\lambda \\
 &= \int \mathbb{E}(\psi|\mathbf{KY}) \otimes \mathbb{E}(\phi|\mathbf{KZ}) \, d\lambda = \int \psi \, d\mu_{\mathbf{Y}} \int \phi \, d\mu_{\mathbf{Z}},
 \end{aligned}$$

where we have used, in the last equality, the fact that  $\lambda$  projects to the product joining of  $\mathbf{KY}$  and  $\mathbf{KZ}$ . This establishes (8) and, therefore, the claim.  $\square$

### 4.3. Quasi-disjointness is preserved by inverse limits.

**THEOREM 4.8.** *Let  $\mathbf{X}$  and  $\mathbf{Z}$  be  $G$  systems and assume that  $\mathbf{X}$  is the inverse limit of a sequence  $n \mapsto \mathbf{X}_n$  of  $G$  systems. If each  $\mathbf{X}_n$  is quasi-disjoint from  $\mathbf{Z}$ , then so is  $\mathbf{X}$ .*

*Proof.* Fix a joining  $\lambda$  of  $\mathbf{X}$  and  $\mathbf{Z}$  whose projection to a joining of  $\mathbf{KX}$  with  $\mathbf{KZ}$  is the product measure. For every  $n$ , the system  $\mathbf{KX}_n$  is a factor of  $\mathbf{KX}$  so  $\lambda$  projects



to the product joining of  $KX_n$  with  $KZ$ . As  $X_n$  and  $Z$  are assumed quasi-disjoint, the projection of  $\lambda$  to a joining of  $X_n$  with  $Z$  is the product measure  $\mu_{X_n} \otimes \mu_Z$ . Therefore,  $\lambda(A \times B) = \mu_X(A)\mu_Z(B)$  for all measurable  $A \subset X_n$  and all  $B \subset Z$  for all  $n \in \mathbb{N}$ . In other words,  $\lambda$  and  $\mu_X \otimes \mu_Z$  agree on all measurable sets of the form  $A \times B$  where  $A \subset X_n$  for some  $n \in \mathbb{N}$  and  $B \subset Z$ . As the  $\sigma$ -algebra generated by such sets is the Borel  $\sigma$ -algebra on  $X \times Z$  we must have  $\lambda = \mu_X \otimes \mu_Z$  as desired.  $\square$

4.4. *An example.* It follows from Theorem 1.1 that any ergodic measurably distal  $\mathbb{Z}$  system is quasi-disjoint from itself. However, the converse is not true.

*Example 4.9.* There exists an ergodic  $\mathbb{Z}$  system  $\mathbf{X}$  which is not measurably distal but is quasi-disjoint from itself and every other ergodic system.

*Proof.* In [GW89, Theorem 2.2], Glasner and Weiss constructed a continuous  $\mathbb{Z}$  action on a compact, metric space  $X$  that is minimal and uniquely ergodic, but for which the corresponding  $\mathbb{Z}$  system  $(X, T, \mu_X)$  is not measurably distal. Moreover, they prove for their system that the only invariant measure on the set  $\tilde{X} = \{\mu \in \mathcal{M}(X) : \pi_{\mathbf{X}, K\mathbf{X}}(\mu) = \mu_{K\mathbf{X}}\}$  is the Dirac measure  $\delta_{\mu_X}$  at the unique invariant measure  $\mu_X \in \tilde{X}$  of  $X$ .

Given any ergodic system  $\mathbf{Y}$ , let  $\lambda \in \mathcal{J}_e(\mathbf{X}, \mathbf{Y})$  and assume that  $\lambda$  projects to the product measure in  $K\mathbf{X} \times K\mathbf{Y}$ . Let  $\lambda = \int_Y \lambda_y d\mu_Y(y)$  be the disintegration of  $\lambda$  with respect to the factor map  $(X \times Y, \lambda) \rightarrow (Y, \mu_Y)$ . Observe that

$$\mu_{K\mathbf{X}} \otimes \mu_{K\mathbf{Y}} = \int_Y \pi_{\mathbf{X}, K\mathbf{X}}(\lambda_y) \otimes \pi_{\mathbf{Y}, K\mathbf{Y}}(\delta_y) d\mu_Y(y)$$

and so  $\pi_{\mathbf{X}, K\mathbf{X}}(\lambda_y) = \mu_{K\mathbf{X}}$ , which implies that  $\lambda_y \in \tilde{X}$ .

Let  $\nu$  be the measure on  $\tilde{X}$  obtained as the pushforward of  $\mu_Y$  by the map  $y \mapsto \lambda_y$ . As  $\nu$  is invariant, it follows that  $\nu = \delta_{\mu_X}$  and, therefore,  $\lambda = \mu_X \otimes \mu_Y$ . We conclude that  $\mathbf{X}$  and  $\mathbf{Y}$  are quasi-disjoint as desired.  $\square$

## 5. Proof of Theorem 1.6

In this section, we prove Theorem 1.6.

*Proof of Theorem 1.6.* In view of a version for amenable groups of the Jewett–Krieger theorem [Ros86], we can assume that  $\mathbf{Y}$  is uniquely ergodic. Fix  $(X, T, \mu_X)$  Kronecker disjoint from  $(Y, S, \mu_Y)$ , a function  $f \in C(X)$ , and a point  $x \in X$ . Kronecker disjointness together with Theorem 1.4 implies that the sequence

$$\frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} \delta_{(T^g x, S^g y)}$$

of measures converges to  $\mu_X \otimes \mu_Y$  for every  $y \in Y$ . This implies that we can take  $Y' = Y$  when  $\phi$  is continuous.

For the general case, fix  $\phi$  in  $L^1(Y, \mu_Y)$ . Let  $k \mapsto \phi_k$  be a sequence in  $C(Y)$  with  $\phi_k \rightarrow \phi$  in  $L^1(Y, \mu_Y)$ . Let  $Y'$  be the set of points  $y \in Y$  such that

$$\lim_{N \rightarrow \infty} \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} |\phi_k(S^g y) - \phi(S^g y)| = \|\phi_k - \phi\|_1 \tag{9}$$

for all  $k \in \mathbb{N}$ . In view of Lindenstrauss' pointwise ergodic theorem [Lin01], we have that  $\mu_Y(Y') = 1$ . Next, let  $y \in Y'$ , let  $f \in C(X)$ ,  $x \in X$ , and let  $k, N \in \mathbb{N}$ . By rescaling, assume that  $\sup\{f(x) : x \in X\} \leq 1$ . We have

$$\begin{aligned} & \left| \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} f(T^g x) \phi_k(S^g y) - \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} f(T^g x) \phi(S^g y) \right| \\ & \leq \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} |\phi_k(S^g y) - \phi(S^g y)|, \end{aligned} \tag{10}$$

$$\lim_{N \rightarrow \infty} \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} f(T^g x) \phi_k(S^g y) = \int_X f \, d\mu_X \int_Y \phi_k \, d\mu_Y \tag{11}$$

and putting (9), (10), and (11) together, we obtain

$$\limsup_{N \rightarrow \infty} \left| \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} f(T^g x) \phi(S^g y) - \int_X f \, d\mu_X \int_Y \phi \, d\mu_Y \right| \leq \|\phi_k - \phi\|_1$$

for every  $k \in \mathbb{N}$ . As  $\phi_k \rightarrow \phi$  in  $L^1$ , we obtain the desired result. □

We now see how to derive the classical Wiener–Wintner theorem from Theorem 1.6.

**COROLLARY 5.1.** (Wiener–Wintner theorem) *Let  $\mathbf{X} = (X, T, \mu_X)$  be a  $\mathbb{Z}$  system and let  $f \in L^1(X, \mu_X)$ . There exists a set  $X_0 \subset X$  with  $\mu(X_0) = 1$  such that for every  $\alpha \in \mathbb{R}$  and every  $x \in X_0$  the limit*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(T^n x) e(n\alpha) \tag{12}$$

*exists.*

*Proof.* Denote by  $\mathbb{T}$  the circle  $\mathbb{R}/\mathbb{Z}$ . For each  $\alpha \in \mathbb{R}$ , let  $R_\alpha : \mathbb{T} \rightarrow \mathbb{T}$  be the rotation  $R_\alpha : t \mapsto t + \alpha$ . The pointwise ergodic theorem of Birkhoff applied to the system  $(X \times \mathbb{T}, T \times R_\alpha)$  implies that there exists a set  $X_\alpha \subset X$  with full measure such that, for every  $x \in X_\alpha$ , the limit (12) exists. The discrete spectrum  $\text{Eig}(\mathbf{X})$  of  $\mathbf{X}$  (cf. Definition 2.2) is at most countable, so the intersection  $X_1 = \bigcap_{\alpha \in \text{Eig}(\mathbf{X})} X_\alpha$  still has full measure.

Next, let  $X_2 \subset X$  be the full measure set given by Theorem 1.6 applied to  $\mathbf{X}$  in place of  $\mathbf{Y}$ . As for every  $\alpha \notin \text{Eig}(\mathbf{X})$  the systems  $\mathbf{X}$  and  $(\mathbb{T}, R_\alpha)$  are Kronecker disjoint, it follows that for every  $x \in X_2$ , the limit in (12) exists. Therefore, the limit exists for every  $\alpha \in \mathbb{R}$  and every  $x \in X_0 = X_1 \cap X_2$ . □

## 6. Applications to multicorrelation sequences and multiple recurrence

In this section, we present applications of our main results to the theory of multiple recurrence. In particular, this section contains proofs of Theorems 1.7 and 1.8. We remark that Theorems 1.7 and 1.8 only concern measure-preserving  $\mathbb{Z}$  systems. The analogs for more general groups  $G$  remain open. We formulate some open questions in this direction in §7.

6.1. *Preliminaries on nilmanifolds.* In [HK05], Host and Kra established a structure theorem for multiple ergodic averages which revealed a deep connection between multi-correlation sequences and single-orbit dynamics on compact nilmanifolds. In the proofs of Theorems 1.7 and 1.8, we make use of refinements of the Host–Kra structure theorem which appeared in [BHK05, HK09, MR19, MR21]. The purpose of this subsection is to give an overview of these results and some related methods that we use in the subsequent sections.

We begin with the definition of a nilmanifold. A closed subgroup  $\Gamma$  of  $G$  is called *uniform* if  $G/\Gamma$  is compact or, equivalently, if there exists a compact set  $K$  such that  $K\Gamma = G$ . Let  $G$  be a  $k$ -step nilpotent Lie group and let  $\Gamma \subset G$  be a uniform and discrete subgroup of  $G$ . The quotient space  $G/\Gamma$  is called a *nilmanifold*. Naturally,  $G$  acts continuously and transitively on  $G/\Gamma$  via left-multiplication, that is,  $a(g\Gamma) = (ag)\Gamma$  for all  $a \in G$  and all  $g\Gamma \in G/\Gamma$ . On any nilmanifold  $G/\Gamma$ , there exists a unique  $G$  invariant Borel probability measure called the *Haar measure of  $G/\Gamma$* , which we denote by  $\mu_{G/\Gamma}$  (cf. [Rag72, Lemma 1.4]). Given a fixed group element  $a \in G$  the map  $R : G/\Gamma \rightarrow G/\Gamma$  defined by  $R(x) = ax$  for all  $x = g\Gamma \in G/\Gamma$  is a *niltranslation* and the resulting  $\mathbb{Z}$  system  $(G/\Gamma, R, \mu_{G/\Gamma})$  is called a  *$k$ -step nilsystem*. We remark that  $(G/\Gamma, R, \mu_{G/\Gamma})$  is ergodic if and only if  $R$  acts transitively on  $G/\Gamma$ ; in fact, if  $R$  is transitive, then  $\mu_{G/\Gamma}$  is the unique  $R$  invariant Borel probability measure on  $G/\Gamma$  (cf. [AGH63, Par69]). Finally, for any  $x \in G/\Gamma$ , the orbit closure  $Y = \overline{\{R^n x : n \in \mathbb{Z}\}} \subset G/\Gamma$  is a *sub-nilmanifold* of  $G/\Gamma$ , meaning that there exists a closed subgroup  $H$  of  $G$  such that  $a \in H$ ,  $Y = Hx$  and  $\Lambda = H \cap \Gamma$  is a uniform and discrete subgroup of  $H$ . In this case, there exists a unique  $H$  invariant Borel probability measure  $\mu_Y$  on  $Y$ , called the Haar measure of the sub-nilmanifold  $Y$ , and the system  $(Y, R, \mu_Y)$  is a nilsystem, as it is isomorphic to  $(H/\Lambda, R, \mu_{H/\Lambda})$  (cf. [Lei06]). For more information on nilmanifolds and nilsystems, we refer the reader to [AGH63, Par69, Par70, Rag72].

**THEOREM 6.1.** (See [MR21, Revised Theorem 7.1]) *Let  $k \in \mathbb{N}$ , let  $X = (G/\Gamma, R, \mu_{G/\Gamma})$  be an ergodic nilrotation and assume that  $G/\Gamma$  is connected. Define  $S = R \times R^2 \times \cdots \times R^k$  and for every  $x \in G/\Gamma$  consider the sub-nilmanifold  $\Omega(X, x)$  of  $(G/\Gamma)^k = G^k/\Gamma^k$  defined as*

$$\Omega(X, x) = \overline{\{S^n(x, x, \dots, x) : n \in \mathbb{Z}\}}.$$

*Let  $\mu_{\Omega(X, x)}$  denote the Haar measure on  $\Omega(X, x)$  and let  $\theta \in [0, 1)$ . If  $e(\theta) \notin \text{Eig}(X)$ , then for almost every  $x \in G/\Gamma$ ,  $e(\theta) \notin \text{Eig}(\Omega(X, x), S, \mu_{\Omega(X, x)})$ .*

For our purposes, we need a generalization of Theorem 6.1 that holds for nilmanifolds  $G/\Gamma$  that are not necessarily connected.

**THEOREM 6.2.** *Let  $k \in \mathbb{N}$  and let  $\mathbf{Z} = (G/\Gamma, R, \mu_{G/\Gamma})$  be an ergodic nilsystem. Define  $S = R \times R^2 \times \dots \times R^k$  and*

$$\Omega(\mathbf{Z}, x) = \overline{\{S^n(x, x, \dots, x) : n \in \mathbb{Z}\}} \subset (G/\Gamma)^k.$$

*For any  $\theta \in [0, 1)$ , if  $e(\theta) \notin \text{Eig}(\mathbf{Z})$  then for almost every  $x \in G/\Gamma$ ,  $e(\theta) \notin \text{Eig}(\Omega(\mathbf{Z}, x), S, \mu_{\Omega(\mathbf{Z}, x)})$ .*

To derive Theorem 6.2 from Theorem 6.1, we need the following well-known lemma regarding nilsystems.

**LEMMA 6.3.** (Cf. [AGH63, Lei05, Par69]) *Suppose  $\mathbf{Z} = (G/\Gamma, R, \mu_{G/\Gamma})$  is a nilsystem. Then the following are equivalent:*

- (i)  $G/\Gamma$  is connected and  $\mathbf{Z}$  is ergodic;
- (ii)  $\mathbf{Z}$  is totally ergodic.

*Proof of Theorem 6.2.* Suppose  $G/\Gamma$  is not connected. As  $G/\Gamma$  is compact, it splits into finitely many distinct connected components  $Z_0, Z_1, \dots, Z_{t-1}$ . It is also straightforward to show that  $Z_i$  is itself a nilmanifold with Haar measure  $\mu_{Z_i}$  and that  $\mu_{G/\Gamma} = 1/t(\mu_{Z_0} + \mu_{Z_1} + \dots + \mu_{Z_{t-1}})$ . The ergodic niltranslation  $R : G/\Gamma \rightarrow G/\Gamma$  cyclically permutes these connected components, so after re-indexing them, if necessary, we have  $R^{tn+i} Z_0 = Z_i$  for all  $n \in \mathbb{N}$  and  $i \in \{0, 1, \dots, t-1\}$ . In particular, for every  $i \in \{0, \dots, t-1\}$  the component  $Z_i$  is  $R^t$  invariant and  $R^t : Z_i \rightarrow Z_i$  is an ergodic niltranslation on  $Z_i$ . Let  $\mathbf{Z}_i = (Z_i, R^t, \mu_{Z_i})$ .

As  $Z_i$  is connected, it follows from Lemma 6.3 that  $\mathbf{Z}_i$  is totally ergodic. This means that  $\text{Eig}(\mathbf{Z}_i)$  contains no roots of unity. Also note that the function  $\sum_{i=0}^{t-1} e(i/t)1_{Z_i}$  is an eigenfunction for  $R$  with eigenvalue  $e(i/t)$ , where  $e(x) = e^{2\pi ix}$  for all  $x \in \mathbb{R}$ . We conclude that

$$\text{Eig}(\mathbf{Z}) \cap \{\text{roots of unity}\} = \left\{ 1, e\left(\frac{1}{t}\right), e\left(\frac{2}{t}\right), \dots, e\left(\frac{t-1}{t}\right) \right\}.$$

On the other hand, if  $\zeta$  is not a root of unity, then  $\zeta$  is an eigenvalue for  $R$  if and only if  $\zeta^t$  is an eigenvalue for  $R^t$ ; therefore,

$$\text{Eig}(\mathbf{Z}) = (\text{Eig}(\mathbf{Z}_0))^{1/t} \cdot \left\{ 1, e\left(\frac{1}{t}\right), e\left(\frac{2}{t}\right), \dots, e\left(\frac{t-1}{t}\right) \right\}.$$

Let  $x \in Z$  and let  $i_0 \in \{0, 1, \dots, t-1\}$  be such that  $x \in Z_{i_0}$ . Define

$$\Omega(\mathbf{Z}_{i_0}, x) = \overline{\{S^{tn}(x, x, \dots, x) : n \in \mathbb{Z}\}}.$$

Observe that for  $i \in \{0, \dots, t-1\}$ ,

$$S^i \Omega(\mathbf{Z}_{i_0}, x) \subset Z_{i_0+i} \times Z_{i_0+2i} \times \dots \times Z_{i_0+ki}$$

and, hence,  $S^i \Omega(\mathbf{Z}_{i_0}, x) \cap S^j \Omega(\mathbf{Z}_{i_0}, x) = \emptyset$  for  $i \neq j$ . As  $\mathbf{Z}_{i_0}$  is totally ergodic, it follows from Theorem 6.1 that for almost every  $x \in Z_{i_0}$ , the nilsystem  $(\Omega(\mathbf{Z}_{i_0}, x), S^t)$  is also

totally ergodic. In view of Lemma 6.3, this means that  $\Omega(\mathbf{Z}_{i_0}, x)$  is connected. We deduce that for almost every  $x \in Z$ , the nilmanifold  $\Omega(\mathbf{Z}, x)$  has  $t$  connected components, because

$$\Omega(\mathbf{Z}, x) = \bigcup_{i=0}^{t-1} S^i \Omega(\mathbf{Z}_{i_0}, x),$$

where  $\Omega(\mathbf{Z}_{i_0}, x), S^1 \Omega(\mathbf{Z}_{i_0}, x), \dots, S^{t-1} \Omega(\mathbf{Z}_{i_0}, x)$  are connected and distinct.

Finally, suppose  $\theta \in [0, 1)$  is such that  $e(\theta) \notin \text{Eig}(\mathbf{Z})$ . Then  $e(\theta)^i \notin \text{Eig}(\mathbf{Z}_i)$  for all  $i$ . From Theorem 6.1, we deduce that for almost every  $x \in Z$ ,  $e(\theta)^i \notin \text{Eig}(\Omega(\mathbf{Z}_{i_0}, x), S^i, \mu_{\Omega(\mathbf{Z}_{i_0}, x)})$  (where, again,  $i_0 \in \{0, \dots, t-1\}$  is such that  $x \in \mathbf{Z}_{i_0}$ ) and, hence,  $e(\theta) \notin \text{Eig}(\Omega(\mathbf{Z}, x), S, \mu_{\Omega(\mathbf{Z}, x)})$ .  $\square$

**COROLLARY 6.4.** *Let  $k \in \mathbb{N}$  and let  $\mathbf{Z} = (G/\Gamma, R, \mu_{G/\Gamma})$  be an ergodic nilsystem. Define  $S = R \times R^2 \times \dots \times R^k$  and*

$$\Omega(\mathbf{Z}, x) = \overline{\{S^n(x, x, \dots, x) : n \in \mathbb{Z}\}} \subset (G/\Gamma)^k.$$

*Let  $\mathbf{X} = (X, T, \mu_X)$  be a  $\mathbb{Z}$  system. If  $\mathbf{X}$  and  $\mathbf{Z}$  are Kronecker disjoint, then for almost every  $x \in G/\Gamma$ , the  $\mathbb{Z}$  systems  $\mathbf{X}$  and  $(\Omega(\mathbf{Z}, x), S, \mu_{\Omega(\mathbf{Z}, x)})$  are Kronecker disjoint.*

*Proof.* By way of contradiction, assume that there is a positive measure set  $X' \subset G/\Gamma$  such that  $\mathbf{X}$  and  $(\Omega(\mathbf{Z}, x), S, \mu_{\Omega(\mathbf{Z}, x)})$  are not Kronecker disjoint whenever  $x \in X'$ . This means that for any  $x \in X'$ , we can find some  $\theta_x \in [0, 1)$  such that  $e(\theta_x)$  is a common eigenvalue for the systems  $\mathbf{X}$  and  $(\Omega(\mathbf{Z}, x), S, \mu_{\Omega(\mathbf{Z}, x)})$ . As  $\mathbf{X}$  only possesses countably many eigenvalues, there exists a positive measure subset  $X'' \subset X'$  such that  $\theta_x = \theta$  is constant for all  $x \in X''$ . As  $e(\theta)$  belongs to  $\text{Eig}(\Omega(\mathbf{Z}, x), S, \mu_{\Omega(\mathbf{Z}, x)})$  for all  $x \in X''$  and  $X''$  has positive measure, it follows from Theorem 6.2 that  $e(\theta)$  belongs to  $\text{Eig}(\mathbf{Z})$ . This contradicts the assumption that  $\mathbf{X}$  and  $\mathbf{Z}$  are Kronecker disjoint.  $\square$

**6.2. Proofs of Theorems 1.7 and 1.8.** The following result, which is used in the proofs of Theorems 1.7 and 1.8, is contained implicitly in [HK09, §7.3].

**THEOREM 6.5.** (Cf. [HK09, §7.3]) *Let  $k \in \mathbb{N}$ , let  $\mathbf{X} = (X, T, \mu_X)$  be a  $\mathbb{Z}$  system, and let  $f_1, \dots, f_k \in L^\infty(X, \mu_X)$ . Then, for every  $\varepsilon > 0$ , there exists a  $k$ -step nilsystem  $(G/\Gamma, R, \mu_{G/\Gamma})$ , which is a factor of  $(X, T, \mu_X)$ , and there exist continuous functions  $g_1, \dots, g_k \in C(G/\Gamma)$  such that for every bounded complex-valued sequence  $(a_n)_{n \in \mathbb{N}}$ , one has*

$$\limsup_{N \rightarrow \infty} \left\| \frac{1}{N} \sum_{n=1}^N a_n \prod_{i=0}^k T^{in} f_i - \frac{1}{N} \sum_{n=1}^N a_n \prod_{i=0}^k (R^{in} g_i) \circ \pi \right\|_2 \leq \varepsilon \sup_{n \in \mathbb{N}} |a_n|,$$

where  $\pi : X \rightarrow G/\Gamma$  denotes the factor map from  $(X, T, \mu_X)$  onto  $(G/\Gamma, R, \mu_{G/\Gamma})$ .

We also need the following lemma.

**LEMMA 6.6.** *Let  $(Y, S)$  be a topological  $\mathbb{Z}$  system, let  $\mu_Y$  be an ergodic  $S$  invariant Borel probability measure on  $Y$ , and let  $G \in L^1(Y, \mu_Y)$ . Then there exists a set  $Y' \subset Y$  with  $\mu_Y(Y') = 1$  such that for any ergodic nilsystem system  $(G/\Gamma, \mu_{G/\Gamma}, R)$  which is*

Kronecker disjoint from  $(Y, S, \mu_Y)$ , any  $F \in C(G/\Gamma)$ , any  $x \in G/\Gamma$ , and any  $y \in Y'$  we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N F(R^n x)G(S^n y) = \int_{G/\Gamma} F d\mu_{G/\Gamma} \int_Y G d\mu_Y.$$

Moreover, if  $(Y, S)$  is uniquely ergodic and  $g \in C(Y)$ , then we can take  $Y' = Y$ .

*Proof.* As any nilsystem is measurably distal (see [Lei05, Theorem 2.14]), Lemma 6.6 is a special case of Theorem 1.6. □

*Proof of Theorem 1.7.* Let  $\mathbf{X} = (X, T, \mu_X)$  and  $\mathbf{Y} = (Y, S, \mu_Y)$  be  $\mathbb{Z}$  systems and assume  $\mathbf{X}$  and  $\mathbf{Y}$  are Kronecker disjoint. Let  $k, \ell \in \mathbb{N}$ ,  $f_1, \dots, f_k \in L^\infty(X, \mu_X)$  and  $g_1, \dots, g_\ell \in L^\infty(Y, \mu_Y)$ . According to [HK05, Zie07], the limit

$$F = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k T^{in} f_i$$

exists in  $L^2(X, \mu_X)$  and the limit

$$G = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \prod_{j=1}^\ell S^{jn} g_j$$

exists in  $L^2(Y, \mu_Y)$ . Moreover, in view of [Tao08] also the limit

$$H = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k \prod_{j=1}^\ell T^{in} f_i S^{jn} g_j$$

exists in  $L^2(X \times Y, \mu_X \otimes \mu_Y)$ . Our goal is to show that

$$H = F \otimes G. \tag{13}$$

We can assume without loss of generality that  $\|f_i\|_\infty \leq 1$  and  $\|g_j\|_\infty \leq 1$ . Fix  $\varepsilon > 0$ . First, we apply Theorem 6.5 to find a  $k$ -step nilsystem  $(G_X/\Gamma_X, R_X, \mu_{G_X/\Gamma_X})$ , which is a factor of  $(X, T, \mu_X)$ , and a set of continuous functions  $\tilde{f}_1, \dots, \tilde{f}_k \in C(G_X/\Gamma_X)$  such that for every bounded complex-valued sequence  $(a_n)_{n \in \mathbb{N}}$  one has

$$\limsup_{N \rightarrow \infty} \left\| \frac{1}{N} \sum_{n=1}^N a_n \prod_{i=1}^k T^{in} f_i - \frac{1}{N} \sum_{n=1}^N a_n \prod_{i=1}^k (R_X^{in} \tilde{f}_i) \circ \pi \right\|_2 \leq \varepsilon \sup_{n \in \mathbb{N}} |a_n|,$$

where  $\pi : X \rightarrow G_X/\Gamma_X$  denotes the factor map from  $(X, T, \mu_X)$  onto  $(G_X/\Gamma_X, R_X, \mu_{G_X/\Gamma_X})$ . In particular, if we choose  $a_n = \prod_{j=1}^\ell S^{jn} g_j(y)$  as  $y$  runs through  $Y$ , we obtain

$$\begin{aligned} & \sup_{y \in Y} \limsup_{N \rightarrow \infty} \left\| \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k \prod_{j=1}^\ell T^{in} f_i S^{jn} g_j(y) \right. \\ & \left. - \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k \prod_{j=1}^\ell ((R_X^{in} \tilde{f}_i) \circ \pi) S^{jn} g_j(y) \right\|_2 \leq \varepsilon. \end{aligned} \tag{14}$$

From (14), it follows that

$$\limsup_{N \rightarrow \infty} \left\| \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k \prod_{j=1}^{\ell} T^{in} f_i S^{jn} g_j - \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k \prod_{j=1}^{\ell} ((R_X^{in} \tilde{f}_i) \circ \pi) S^{jn} g_j \right\|_2 \leq \varepsilon. \tag{15}$$

Similarly, we can pick a  $\ell$ -step nilsystem  $(G_Y/\Gamma_Y, R_Y, \mu_{G_Y/\Gamma_Y})$ , which is a factor of  $(Y, S, \mu_Y)$ , and a set of continuous functions  $\tilde{g}_1, \dots, \tilde{g}_\ell \in C(G_Y/\Gamma_Y)$  such that, for every bounded complex-valued sequence  $(b_n)_{n \in \mathbb{N}}$ , one has

$$\limsup_{N \rightarrow \infty} \left\| \frac{1}{N} \sum_{n=1}^N b_n \prod_{j=1}^{\ell} S^{jn} g_j - \frac{1}{N} \sum_{n=1}^N b_n \prod_{j=1}^{\ell} (R_Y^{jn} \tilde{g}_j) \circ \eta \right\|_2 \leq \varepsilon \sup_{n \in \mathbb{N}} |b_n|,$$

where  $\eta : Y \rightarrow G_Y/\Gamma_Y$  denotes the factor map from  $(Y, S, \mu_Y)$  onto  $(G_Y/\Gamma_Y, R_Y, \mu_{G_Y/\Gamma_Y})$ . If we set  $b_n = \prod_{i=1}^k (R_X^{in} \tilde{f}_i) \circ \pi(x)$ , then we obtain

$$\begin{aligned} \sup_{x \in X} \limsup_{N \rightarrow \infty} & \left\| \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k \prod_{j=1}^{\ell} (R_X^{in} \tilde{f}_i)(\pi x) S^{jn} g_j \right. \\ & \left. - \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k \prod_{j=1}^{\ell} (R_X^{in} \tilde{f}_i)(\pi x) ((R_Y^{jn} \tilde{g}_j) \circ \eta) \right\|_2 \leq \varepsilon \end{aligned} \tag{16}$$

and, hence,

$$\begin{aligned} \limsup_{N \rightarrow \infty} & \left\| \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k \prod_{j=1}^{\ell} ((R_X^{in} \tilde{f}_i) \circ \pi) S^{jn} g_j \right. \\ & \left. - \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k \prod_{j=1}^{\ell} ((R_X^{in} \tilde{f}_i) \circ \pi) ((R_Y^{jn} \tilde{g}_j) \circ \eta) \right\|_2 \leq \varepsilon. \end{aligned} \tag{17}$$

Combining (15) and (17) yields

$$\limsup_{N \rightarrow \infty} \left\| \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k \prod_{j=1}^{\ell} T^{in} f_i S^{jn} g_j - \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k \prod_{j=1}^{\ell} ((R_X^{in} \tilde{f}_i) \circ \pi) ((R_Y^{jn} \tilde{g}_j) \circ \eta) \right\|_2 \leq 2\varepsilon. \tag{18}$$

Next, we claim that for almost every  $x \in G_X/\Gamma_X$  and almost every  $y \in G_Y/\Gamma_Y$ , we have

$$\begin{aligned} \lim_{N \rightarrow \infty} & \left| \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k \prod_{j=1}^{\ell} \tilde{f}_i(R_X^{in} x) \tilde{g}_j(R_Y^{jn} y) \right. \\ & \left. - \left( \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k \tilde{f}_i(R_X^{in} x) \right) \left( \frac{1}{N} \sum_{n=1}^N \prod_{j=1}^{\ell} \tilde{g}_j(R_Y^{jn} y) \right) \right| = 0. \end{aligned} \tag{19}$$



Assume for now that this claim holds. It follows from (14) and (16) that

$$\begin{aligned} \limsup_{N \rightarrow \infty} & \left\| \left( \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k R_X^{in} \tilde{f}_i \circ \pi \right) \left( \frac{1}{N} \sum_{n=1}^N \prod_{j=1}^\ell R_Y^{jn} \tilde{g}_j \circ \eta \right) \right. \\ & \left. - \left( \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k T^{in} f_i \right) \left( \frac{1}{N} \sum_{n=1}^N \prod_{j=1}^\ell S^{jn} g_j \right) \right\|_2 \leq 2\varepsilon. \end{aligned} \tag{20}$$

Thus, combining (20) with (19) and (18) gives

$$\lim_{N \rightarrow \infty} \left\| \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k \prod_{j=1}^\ell T^{in} f_i S^{jn} g_j - \left( \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k T^{in} f_i \right) \cdot \left( \frac{1}{N} \sum_{n=1}^N \prod_{j=1}^\ell S^{jn} g_j \right) \right\|_2 \leq 4\varepsilon.$$

As  $\varepsilon > 0$  was chosen arbitrarily, the proof of (13) is complete.

It remains to show that (19) is true. Define  $S_X = R_X \times R_X^2 \times \cdots \times R_X^k$  and for every  $x \in G_X/\Gamma_X$  consider  $\Omega(\mathbf{X}, x) = \{S_X^n(x, x, \dots, x) : n \in \mathbb{Z}\}$ . In addition, define  $F = \tilde{f}_1 \otimes \cdots \otimes \tilde{f}_k$ . Similarly, we define  $S_Y = R_Y \times R_Y^2 \times \cdots \times R_Y^\ell$ ,  $\Omega(\mathbf{Y}, y) = \{S_Y^n(y, y, \dots, y) : n \in \mathbb{Z}\}$  and  $G = \tilde{g}_1 \otimes \cdots \otimes \tilde{g}_\ell$ . Hence, (19) can be rewritten as

$$\begin{aligned} \lim_{N \rightarrow \infty} & \left| \frac{1}{N} \sum_{n=1}^N F(S_X^n(x, \dots, x)) G(S_Y^n(y, \dots, y)) \right. \\ & \left. - \left( \frac{1}{N} \sum_{n=1}^N F(S_X^n(x, \dots, x)) \right) \left( \frac{1}{N} \sum_{n=1}^N G(S_Y^n(y, \dots, y)) \right) \right| = 0. \end{aligned} \tag{21}$$

Note that  $(G_X/\Gamma_X, R_X, \mu_{G_X/\Gamma_X})$  and  $(G_Y/\Gamma_Y, R_Y, \mu_{G_Y/\Gamma_Y})$  are Kronecker disjoint, because  $(X, T, \mu_X)$  and  $(Y, S, \mu_Y)$  are Kronecker disjoint. In view of Corollary 6.4 it therefore follows that for  $\mu_X$ -almost every  $x \in G_X/\Gamma_X$  and for  $\mu_Y$ -almost every  $y \in G_Y/\Gamma_Y$  the two nilsystems  $(\Omega(\mathbf{X}, x), S_X, \mu_{\Omega(\mathbf{X},x)})$  and  $(\Omega(\mathbf{Y}, y), S_Y, \mu_{\Omega(\mathbf{Y},y)})$  are Kronecker disjoint. We can now apply Lemma 6.6 to conclude that for almost every  $x \in G_X/\Gamma_X$  and almost every  $y \in G_Y/\Gamma_Y$ , we have

$$\begin{aligned} \lim_{N \rightarrow \infty} & \frac{1}{N} \sum_{n=1}^N F(S_X^n(x, \dots, x)) G(S_Y^n(y, \dots, y)) \\ & = \int_{\Omega(\mathbf{X},x)} F \, d\mu_{\Omega(\mathbf{X},x)} \int_{\Omega(\mathbf{Y},y)} G \, d\mu_{\Omega(\mathbf{Y},y)}. \end{aligned} \tag{22}$$

As  $(\Omega(\mathbf{X}, x), S_X)$  and  $(\Omega(\mathbf{Y}, y), S_Y)$  are uniquely ergodic, we have that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N F(S_X^n(x, \dots, x)) = \int_{\Omega(\mathbf{X},x)} F \, d\mu_{\Omega(\mathbf{X},x)}, \tag{23}$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N G(S_Y^n(x, \dots, x)) = \int_{\Omega(\mathbf{Y},y)} G \, d\mu_{\Omega(\mathbf{Y},y)}. \tag{24}$$

Combining (22) with (23) and (24) yields (21), which, in turn, implies (19). This completes the proof of Theorem 1.7. □

*Proof of Theorem 1.8.* Let  $(Y, S)$  be a topological  $\mathbb{Z}$  system, let  $\mu_Y$  be an  $S$  invariant Borel probability measure on  $Y$ , and let  $G \in L^1(Y, \mu_Y)$ . First, we apply Lemma 6.6 to find a set  $Y' \subset Y$  with  $\mu_Y(Y') = 1$  such that for any ergodic nilsystem system  $(G/\Gamma, \mu_{G/\Gamma}, R)$  which is Kronecker disjoint from  $(Y, S, \mu_Y)$ , any  $F \in C(G/\Gamma)$ , any  $x \in G/\Gamma$ , and any  $y \in Y'$  we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N F(x)G(y) = \int_{G/\Gamma} F d\mu_{G/\Gamma} \int_Y G d\mu_Y.$$

Lemma 6.6 also guarantees that if  $(Y, S)$  is uniquely ergodic and  $G \in C(Y)$ , then we can take  $Y' = Y$ . Now let  $(X, T, \mu_X)$  be a  $\mathbb{Z}$  system that is Kronecker disjoint from  $(Y, S, \mu_Y)$ . Fix  $k \in \mathbb{N}$  and let  $f_1, \dots, f_k \in L^\infty(X, \mu_X)$ . Our goal is to show that for any  $y \in Y'$  we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N G(S^n y) \prod_{i=1}^k T^{in} f_i = \left( \int_Y G d\mu_Y \right) \cdot \left( \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k T^{in} f_i \right) \tag{25}$$

in  $L^2(X, \mu_X)$ .

Note that (25) is trivially true if  $G$  is a constant function. Hence, by replacing  $G$  with  $G - \int_Y G d\mu_Y$  if necessary, we can assume without loss of generality that  $\int_Y G d\mu_Y = 0$ . In this case, (25) reduces to

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N G(S^n y) \prod_{i=1}^k T^{in} f_i = 0. \tag{26}$$

We can also assume without loss of generality that  $\|G\|_\infty \leq 1$  and that  $\|f_i\|_\infty \leq 1$  for all  $i = 1, \dots, k$ .

Fix  $\varepsilon > 0$ . We apply Theorem 6.5 to find a  $k$ -step nilsystem  $(G/\Gamma, R, \mu_{G/\Gamma})$ , which is a factor of  $(X, T, \mu_X)$ , and a set of continuous functions  $\tilde{f}_1, \dots, \tilde{f}_k \in C(G/\Gamma)$  such that

$$\limsup_{N \rightarrow \infty} \left\| \frac{1}{N} \sum_{n=1}^N G(S^n y) \prod_{i=1}^k T^{in} f_i - \frac{1}{N} \sum_{n=1}^N G(S^n y) \prod_{i=1}^k (R^{in} \tilde{f}_i) \circ \pi \right\|_2 \leq \varepsilon,$$

where  $\pi : X \rightarrow G/\Gamma$  denotes the factor map from  $(X, T, \mu_X)$  onto  $(G/\Gamma, R, \mu_{G/\Gamma})$ . Therefore, to show (26) it suffices to show that for almost every  $x \in G/\Gamma$  one has

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N G(S^n y) \prod_{i=1}^k R^{in} \tilde{f}_i(x) = 0. \tag{27}$$

Define  $S = R \times R^2 \times \dots \times R^k$ ,  $Y_x = \overline{\{S^n(x, x, \dots, x) : n \in \mathbb{Z}\}}$  and also  $F = \tilde{f}_1 \otimes \dots \otimes \tilde{f}_k$ . Clearly, (27) is equivalent to

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N G(S^n y) S^n F(x, \dots, x) = 0. \tag{28}$$

It follows from Corollary 6.4 that for almost every  $x \in G_X/\Gamma_X$ , the two systems  $(Y_x, S, \mu_{Y_x}, S)$  and  $(Y, S, \mu_Y)$  are Kronecker disjoint. Hence, Lemma 6.6 implies that for almost every  $x \in G_X/\Gamma_X$ , (28) holds. This finishes the proof.  $\square$

6.3. *An example of multiple recurrence.* In this section, we obtain an application of Theorem 1.4 to multiple recurrence.

*Definition 6.7.* Let  $q$  be an integer  $\geq 2$ . A function  $w : \mathbb{N} \rightarrow \mathbb{C}$  is called *strongly  $q$ -multiplicative* if  $w(n) = w(a_0) \cdots w(a_k)$ , where

$$n = \sum_{i=0}^k a_i q^i, \quad 0 \leq a_i \leq q - 1$$

is the base  $q$  expansion of  $n$ . For convenience, we assume  $w(0) = 1$ .

The  $\{-1, 1\}$ -valued Thue–Morse sequence is an example of a strongly 2-multiplicative sequence, obtained by letting  $w(0) = 1$  and  $w(1) = -1$ .

When a strongly  $q$ -multiplicative function takes only finitely many values, say  $w : \mathbb{N} \rightarrow \mathcal{A} \subset \mathbb{C}$ , we can identify the function with a point (which, by a slight abuse of notation, we also denote by  $w$ ) in the symbolic space  $\mathcal{A}^{\mathbb{N}}$ . Let  $S : \mathcal{A}^{\mathbb{N}} \rightarrow \mathcal{A}^{\mathbb{N}}$  be the usual shift map.

**PROPOSITION 6.8.** [Lia87] *Let  $w : \mathbb{N} \rightarrow \mathcal{A} \subset \mathbb{C}$  be a strongly  $q$ -multiplicative function. Then the orbit closure  $Y = (\{S^n w : n \in \mathbb{N}\}, S)$  is a uniquely ergodic system. Moreover, the discrete spectrum  $\text{Eig}(Y)$  with respect to the unique invariant measure  $\mu_Y$  is contained in the set  $\{e(a/q^n) : a, n \in \mathbb{N}\}$ .*

Putting this proposition together with Theorem 1.8, we obtain the following corollary.

**COROLLARY 6.9.** *Let  $w : \mathbb{N} \rightarrow \mathbb{C}$  be a strongly  $q$ -multiplicative function taking only finite many values and let  $X = (X, T, \mu_X)$  be a system with  $\text{Eig}(X) \cap \{e(a/q^n) : a, n \in \mathbb{N}\} = \{1\}$ . For every  $f_1, \dots, f_k \in L^\infty(X)$ , we have*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N w(n) \prod_{i=1}^k T^{in} f_i = \left( \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N w(n) \right) \cdot \left( \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k T^{in} f_i \right).$$

We can use this to derive a multiple recurrence result for level sets of strongly  $q$ -multiplicative functions.

**THEOREM 6.10.** *Let  $m, q \in \mathbb{N}$ , let  $\mathcal{A} \subset \mathbb{C}$  be the set of  $m$ th roots of 1, and let  $w : \mathbb{N} \rightarrow \mathcal{A}$  be a strongly  $q$ -multiplicative function. For every  $z \in \mathcal{A}$ , the level set  $R = \{n \in \mathbb{N} : w(n) = z\}$  either has 0 density or satisfies the following multiple recurrence property: let  $X = (X, T, \mu_X)$  be a  $\mathbb{Z}$  system and assume that  $\text{Eig}(X)$  contains no non-trivial  $q$ th root of 1. Then, for every  $A \subset X$  with  $\mu(A) > 0$  and every  $k \in \mathbb{N}$ , there exists  $n \in R$  such that*

$$\mu(A \cap T^{-n} A \cap T^{-2n} A \cap \cdots \cap T^{-kn} A) > 0.$$

*Proof.* The indicator function  $1_R(n)$  of  $R = \{n \in \mathbb{N} : w(n) = z\}$  can be expressed as  $\delta_z \circ w(n)$ , where  $\delta_z : \mathcal{A} \rightarrow \mathbb{C}$  is the function  $\delta_z(u) = 1$  if  $u = z$  and  $\delta_z(u) = 0$  otherwise. The space of functions from  $\mathcal{A}$  to  $\mathbb{C}$  is a vector space spanned by the functions  $u \mapsto u^j$ ,  $j = 0, \dots, m - 1$ , and, in particular,  $\delta_z$  is a linear combination of the functions  $u^j$ . It follows that  $1_R(n)$  is a linear combination of the functions  $w^j(n)$  for  $j = 0, \dots, m - 1$ .

Observe that each power  $w^j$  of  $w$  is a strongly  $q$ -multiplicative function and, thus, Corollary 6.9 applied to the functions  $f_i = 1_A$  implies that

$$\begin{aligned} & \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N w^j(n) \mu \left( \bigcap_{i=0}^k T^{-in} A \right) \\ &= \left( \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N w^j(n) \right) \cdot \left( \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mu \left( \bigcap_{i=0}^k T^{-in} A \right) \right). \end{aligned}$$

By linearity, this implies that

$$\begin{aligned} & \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N 1_R(n) \mu \left( \bigcap_{i=0}^k T^{-in} A \right) \\ &= \left( \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N 1_R(n) \right) \cdot \left( \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mu \left( \bigcap_{i=0}^k T^{-in} A \right) \right). \end{aligned}$$

If  $R$  has positive density, then the first factor on the right-hand side of the previous equation is positive. The fact that the second factor is also positive is the content of Furstenberg’s multiple recurrence theorem [Fur77, Theorem 11.13]. Therefore, the left-hand side has to be positive as well and this implies the desired conclusion.  $\square$

7. Some open questions

It is tempting to define  $\mathbf{X}$  and  $\mathbf{Z}$  to be quasi-disjoint if the natural map from  $\mathcal{J}_e(\mathbf{X}, \mathbf{Z})$  to  $\mathcal{J}_e(K\mathbf{X}, K\mathbf{Z})$  is a bijection. However, [Ber71, Example 2] shows this is, in fact, a strictly stronger notion than quasi-disjointness, which amounts to requiring that the above map be a bijection almost everywhere with respect to the measure on  $\mathcal{J}_e(\mathbf{X}, \mathbf{Z})$  given by the ergodic decomposition of the product measure. In light of this we ask the following question.

*Question 7.1.* Is it true that a system  $\mathbf{X}$  is quasi-disjoint from  $\mathbf{Y}$  if and only if the support of the measure appearing in the ergodic decomposition of  $\mu \times \nu$  equals  $\mathcal{J}_e(\mathbf{X}, \mathbf{Y})$ .

The notion of disjointness can be described in terms of factor maps. We ask whether a similar characterization of quasi-disjointness is possible.

*Question 7.2.* Is it true that  $\mathbf{X}$  and  $\mathbf{Y}$  are quasi-disjoint if and only if any system  $\mathbf{Z}$  which has  $\mathbf{X}, \mathbf{Y}$  and  $K\mathbf{X} \times K\mathbf{Y}$  as factors, also has  $\mathbf{X} \times \mathbf{Y}$  as a factor?

We are also interested in the following potential extensions of our theorem and its applications.

*Question 7.3.* Is a system  $\mathbf{X}$  quasi-disjoint from any ergodic system if and only if it is quasi-disjoint from itself?

We expect the following question, which seeks a generalization of Theorem 1.7 to countable, amenable groups, to be quite difficult.

*Question 7.4.* Fix a countable, amenable group  $G$  with a Følner sequence  $\Phi$ . Let  $\mathbf{X} = (X, T, \mu_X)$  be an ergodic  $G^k$  system and let  $\mathbf{Y} = (Y, S, \mu_Y)$  be an ergodic  $G^\ell$  system. Given  $1 \leq i \leq j \leq k$ , write  $T_{[i,j]}$  for the  $G$  action induced by the inclusion of  $G$  in  $G^k$  diagonally on the coordinates  $i, \dots, j$  and similarly for  $S_{[i,j]}$  with  $1 \leq i \leq j \leq \ell$ . Under what conditions on  $\mathbf{X}$  and  $\mathbf{Y}$  do we have

$$\begin{aligned} & \lim_{N \rightarrow \infty} \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} \prod_{i=1}^k T_{[1,i]}^g f_i \prod_{j=1}^{\ell} S_{[1,j]}^g h_j \\ &= \left( \lim_{N \rightarrow \infty} \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} \prod_{i=1}^k T_{[1,i]}^g f_i \right) \left( \lim_{N \rightarrow \infty} \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} \prod_{j=1}^{\ell} S_{[1,j]}^g h_j \right) \end{aligned}$$

in  $L^2(X \times Y, \mu_X \otimes \mu_Y)$  for all  $f_1, \dots, f_k$  in  $L^\infty(X, \mu_X)$  and all  $h_1, \dots, h_\ell$  in  $L^\infty(Y, \mu_Y)$ ?

*Acknowledgements.* We would like to thank Vitaly Bergelson for helpful comments on an early version of the paper, Eli Glasner for fruitful discussions on disjointness of measurably distal systems, and Pavel Zorin-Kranich for useful remarks on Wiener–Wintner results. We would like to thank the referee for suggesting improvements to the readability of the paper. We would also like to thank the referee of an earlier version of this paper for pointing out errors in the statements of Theorems 6.1 and 6.2 in light of work by Liang and Qiu, which have since been corrected. The first author was supported by the NSF grant DMS-1700147. The third author gratefully acknowledges the support of the NSF via grants DMS-1246989 and DMS-1703597.

## REFERENCES

- [AGH63] L. Auslander, L. Green and F. Hahn. *Flows on Homogeneous Spaces (Annals of Mathematics Studies, 53)*. Princeton University Press, Princeton, NJ, 1963.
- [Ber71] K. Berg. Quasi-disjointness in ergodic theory. *Trans. Amer. Math. Soc.* **162** (1971), 71–87.
- [Ber72] K. Berg. Quasi-disjointness, products and inverse limits. *Math. Syst. Theory* **6** (1972), 123–128.
- [BHK05] V. Bergelson, B. Host and B. Kra. Multiple recurrence and nilsequences. *Invent. Math.* **160**(2) (2005), 261–303.
- [EW11] M. Einsiedler and T. Ward. *Ergodic Theory with a View Towards Number Theory (Graduate Texts in Mathematics, 259)*. Springer, London, 2011.
- [FK91] H. Furstenberg and Y. Katznelson. A density version of the Hales–Jewett theorem. *J. Anal. Math.* **57** (1991), 64–119.
- [Fur67] H. Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Syst. Theory* **1** (1967), 1–49.
- [Fur77] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Anal. Math.* **31** (1977), 204–256.
- [Fur81] H. Furstenberg. *Recurrence in Ergodic Theory and Combinatorial Number Theory*. Princeton University Press, Princeton, NJ, 1981.
- [Gla03] E. Glasner. *Ergodic Theory via Joinings (Mathematical Surveys and Monographs, 101)*. American Mathematical Society, Providence, RI, 2003.
- [GW89] S. Glasner and B. Weiss. Processes disjoint from weak mixing. *Trans. Amer. Math. Soc.* **316**(2) (1989), 689–703.
- [HK05] B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math. (2)* **161**(1) (2005), pp. 397–488.

- [HK09] B. Host and B. Kra. Uniformity seminorms on  $\ell^\infty$  and applications. *J. Anal. Math.* **108** (2009), 219–276.
- [Lei05] A. Leibman. Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold. *Ergod. Th. & Dynam. Sys.* **25**(1) (2005), 201–213.
- [Lei06] A. Leibman. Rational sub-nilmanifolds of a compact nilmanifold. *Ergod. Th. & Dynam. Sys.* **26**(3) (2006), pp. 787–798.
- [Lei10] A. Leibman. Multiple polynomial correlation sequences and nilsequences. *Ergod. Th. & Dynam. Sys.* **30**(3) (2010), 841–854.
- [LG61] K. de Leeuw and I. Glicksberg. Applications of almost periodic compactifications. *Acta Math.* **105** (1961), 63–97.
- [Lia87] P. Liardet. Regularities of distribution. *Compos. Math.* **61**(3) (1987), 267–293.
- [Lin01] E. Lindenstrauss. Pointwise theorems for amenable groups. *Invent. Math.* **146**(2) (2001), 259–295.
- [Lin99] E. Lindenstrauss. Measurable distal and topological distal systems. *Ergod. Th. & Dynam. Sys.* **19**(4) (1999), 1063–1076.
- [Mac64] G. W. Mackey. Ergodic transformation groups with a pure point spectrum. *Illinois J. Math.* **8** (1964), 593–600.
- [MR19] J. Moreira and F. K. Richter. A spectral refinement of the Bergelson–Host–Kra decomposition and new multiple ergodic theorems. *Ergod. Th. & Dynam. Sys.* **39**(4) (2019), 1042–1070.
- [MR21] J. Moreira and F. K. Richter. A spectral refinement of the Bergelson–Host–Kra decomposition and multiple ergodic theorems—Erratum. *Preprint*, 2021, [arXiv:1609.03631](https://arxiv.org/abs/1609.03631).
- [Par68] W. Parry. Zero entropy of distal and related transformations. *Topological Dynamics (Symposium, Colorado State University, Fort Collins, Colorado, 1967)*. Eds. J. Auslander and W. H. Gottschalk. Benjamin, New York, 1968, pp. 383–389.
- [Par69] W. Parry. Ergodic properties of affine transformations and flows on nilmanifolds. *Amer. J. Math.* **91** (1969), 757–771.
- [Par70] W. Parry. Dynamical systems on nilmanifolds. *Bull. Lond. Math. Soc.* **2** (1970), 37–40.
- [Rag72] M. S. Raghunathan. *Discrete Subgroups of Lie Groups (Ergebnisse der Mathematik und ihrer Grenzgebiete, 68)*. Springer-Verlag, New York, 1972.
- [Ros86] A. Rosenthal. Strictly ergodic models and amenable group action. *PhD Thesis*, Paris 6, 1986.
- [Rud79] D. J. Rudolph. An example of a measure preserving map with minimal self-joinings, and applications. *J. Anal. Math.* **35** (1979), 97–122.
- [Tao08] T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. *Ergod. Th. & Dynam. Sys.* **28**(2) (2008), 657–688.
- [WW41] N. Wiener and A. Wintner. Harmonic analysis and ergodic theory. *Amer. J. Math.* **63** (1941), 415–426.
- [Zie07] T. Ziegler. Universal characteristic factors and Furstenberg averages. *J. Amer. Math. Soc.* **20**(1) (2007), 53–97.
- [Zim76a] R. J. Zimmer. Ergodic actions with generalized discrete spectrum. *Illinois J. Math.* **20**(4) (1976), 555–588.
- [Zim76b] R. J. Zimmer. Extensions of ergodic group actions. *Illinois J. Math.* **20**(3) (1976), 373–409.