

# On embeddability of automorphisms into measurable flows from the point of view of self-joining properties

by

Joanna Kułaga-Przymus (Warszawa and Toruń)

**Abstract.** We compare self-joining and embeddability properties. In particular, we prove that a measure preserving flow  $(T_t)_{t \in \mathbb{R}}$  with  $T_1$  ergodic is 2-fold quasi-simple (resp. 2-fold distally simple) if and only if  $T_1$  is 2-fold quasi-simple (resp. 2-fold distally simple). We also show that the Furstenberg–Zimmer decomposition for a flow  $(T_t)_{t \in \mathbb{R}}$  with  $T_1$  ergodic with respect to any flow factor is the same for  $(T_t)_{t \in \mathbb{R}}$  and for  $T_1$ . We give an example of a 2-fold quasi-simple flow disjoint from simple flows and whose time-one map is simple. We describe two classes of flows (flows with minimal self-joining property and flows with the so-called Ratner property) whose time-one maps have unique embeddings into measurable flows. We also give an example of a 2-fold simple flow whose time-one map has more than one embedding.

## Contents

1. Introduction . . . . .	16
1.1. Embeddability . . . . .	16
1.2. Joining properties . . . . .	17
1.3. Motivation and results . . . . .	18
2. Preliminaries . . . . .	21
2.1. Topology, invariant measures, ergodic decomposition . . . . .	22
2.2. Joinings I: Basic definitions and properties . . . . .	23
2.3. Factors and extensions . . . . .	25
2.4. Joinings II: Self-joining properties . . . . .	26
2.5. Distal and isometric extensions: T-compactness . . . . .	27
3. Self-joining properties of $\mathcal{T}$ and $T_1$ , part I . . . . .	27
3.1. Simplicity . . . . .	27
3.2. Joining primeness property . . . . .	29
3.3. Isometric, distal and relatively weakly mixing extensions . . . . .	32
3.4. Quasi- and distal simplicity . . . . .	38

---

2010 *Mathematics Subject Classification*: Primary 37A05, 37A10; Secondary 37A30, 37A35.

*Key words and phrases*: joinings, quasi-simplicity, distal simplicity, embeddability, uniqueness of embedding.

4. (Non-)uniqueness of embedding . . . . .	39
4.1. Uniqueness . . . . .	39
4.2. Non-uniqueness . . . . .	42
5. Self-joining properties of $\mathcal{T}$ and $T_1$ , part II . . . . .	43
5.1. Easy counterexample . . . . .	44
5.2. Advanced counterexample . . . . .	47
Appendix. Uniform distribution . . . . .	62
A.1. Necessary tools . . . . .	62
A.2. Technical results . . . . .	63
References . . . . .	72

## 1. Introduction

**1.1. Embeddability.** The problem of embeddability of automorphisms into measurable flows has been studied in ergodic theory for more than 60 years. Even today, the basic problem—a necessary and sufficient condition for embeddability—remains the most interesting open question in this area. The first articles on this subject dealt with a simpler problem, namely the existence of roots. In [31] Halmos gave a necessary condition for the existence of a square root (the absence of  $-1$  in the spectrum). In the discrete spectrum case this is also sufficient. These results were later generalized in various ways—both in the discrete [42, 5, 39, 61] and quasi-discrete spectrum [30, 50, 51, 52] case. They include necessary and sufficient conditions for the existence of roots and for the existence of embedding into a flow in the discrete and quasi-discrete spectrum cases <sup>(1)</sup>. In particular, Lamperti [42] gave an example of a discrete spectrum automorphism with roots of all orders, which is however not embeddable into a flow.

The problem turned out to be much more challenging in the weakly mixing case. Already in 1956, Halmos [32] ([33]) asked about the existence of square roots of weakly mixing automorphisms, the existence of roots for Bernoulli shifts and the embeddability of Bernoulli shifts into flows. The answers came relatively fast. Chacon constructed a non-mixing automorphism with continuous spectrum having no square root [7] and a bit later one having no roots at all [8]. Since the centralizer of an automorphism embeddable into a flow contains all time- $t$  automorphisms of this flow, a necessary condition for embeddability is that the centralizer of the automorphism is uncountable. Del Junco [16], continuing the study of Chacon type constructions, showed in particular that the centralizer of Chacon’s classical automorphism [9] consists only of its powers, whence it is too small for the automorphism even to admit roots. Further results were provided by Ornstein: he answered positively the remaining two questions of Halmos on the Bernoulli

---

<sup>(1)</sup> Hahn and Parry [30] showed that automorphisms with quasi-discrete, non-discrete spectrum are not embeddable into flows.

embedding [55, 57]. While all previous constructions were not mixing, and Bernoulli shifts have strong mixing properties, it seemed natural that sufficiently strong mixing properties could imply embeddability (in 1966 this problem was still open [5]). This, however, turned out to be false. Ornstein constructed a mixing automorphism [58] and a K-automorphism [60] without square roots.

After that there seems to have been no substantial progress until 2000 when King [37] showed that a generic automorphism admits roots of all orders. Three years later de la Rue and de Sam Lazaro proved even more: that a generic automorphism can be embedded into a flow [15]. This result was later generalized (in various ways) to  $\mathbb{Z}^d$ -actions [76, 77, 71]. In 2004 Stepin and Eremenko [74] showed that typically there are uncountably many pair-wise spectrally non-isomorphic embeddings, thus strengthening the result from [15].

**1.2. Joining properties.** Joinings were introduced by Furstenberg [26] in 1967, and have since been a very fruitful tool for studying the dynamical systems. A joining of two (or more) systems (in our case automorphisms or flows) is a measure invariant under the product action, whose marginals are the invariant measures of the original systems. We deal with self-joinings, i.e. joinings of copies of a fixed system. Depending on the “number of ways” a system can be joined with itself in an ergodic way, one defines several classes of dynamical systems. We will be mostly interested in 2-fold joinings. The self-joining properties of higher order are defined using the so-called PID property [20] <sup>(2)</sup>. We distinguish systems with *minimal self-joining property* (MSJ) (Rudolph [69]), *2-fold simple* systems (Veech [78], del Junco, Rudolph [20]), *2-fold quasi-simple* (2-QS) systems (Ryzhikov, Thouvenot [70]), *2-fold distally simple* (2-DS) systems (del Junco, Lemańczyk [19]). A particular case of 2-QS systems are systems whose 2-fold ergodic self-joinings other than the product measure are finite extensions of the marginal factors. We will denote them by “ $n : 1$ ” for  $n \geq 1$ . For  $n = 1$  this notion is equivalent to 2-fold simplicity. We also consider *joining primeness property* (JP) (Lemańczyk, Parreau, Roy [45]). Unlike the previous classes, systems with the JP property were originally not defined in terms of self-joinings—instead, some restriction was imposed on the type of ergodic joinings with Cartesian products of weakly mixing systems. Precise definitions are given in Section 2.4, and those classes are related as follows:  $\text{MSJ} \subset \text{2-fold simple} \subset \text{2-QS} \subset \text{2-DS} \subset \text{JP}$ .

One more property we will consider is the so-called *R-property*, which was first observed by Ratner [66] for horocycle flows. It describes the behavior

---

<sup>(2)</sup> Danilenko showed [11] that  $\mathcal{T}$  has the PID property if and only if  $T_1$  has this property.

of orbits of nearby distinct points. Roughly speaking, sufficiently long pieces of their orbits are close to each other, up to a time shift. An important ingredient of the definition of the R-property is the restriction on how large this time shift can be. Precise conditions (see [66, 24, 25, 22]) form a variety of combinatorial properties, all yielding the same dynamical consequences in the case of a weakly mixing flow: each of them implies that the flow is “ $n : 1$ ” for some  $n \in \mathbb{N}$  <sup>(3)</sup>. In fact, it implies more: ergodic joinings with other flows, different from the product measure, are also finite extensions of the coordinate factor of the “additional” flow [66] (see also [75]). Apart from the horocycle flows, several classes of flows satisfying the R-property are known: see [24, 25, 35, 22].

### 1.3. Motivation and results

**1.3.1. Between simplicity and JP: 2-QS and 2-DS.** A starting point for our investigations is a result of del Junco and Rudolph [21] who showed that whenever  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  (with  $T_1$  ergodic) is 2-fold simple then  $T_1$  is also 2-fold simple. On the other hand, it is not difficult to prove that given a flow  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  (with  $T_1$  ergodic),  $\mathcal{T}$  has the JP property whenever  $T_1$  has the JP property (see Section 3.2). Analogous results hold for automorphisms and their powers.

Since the 2-QS and 2-DS properties are “intermediate” between simplicity and the JP property, it is reasonable to expect that 2-QS and 2-DS system will share either the features of simple systems or of JP systems, or both. Moreover, it is quite easy to see that if  $T_1$  is 2-fold simple then  $\mathcal{T}$  is 2-QS. This makes it even less surprising that indeed either both  $\mathcal{T}$  and  $T_1$  are 2-QS (resp. 2-DS) or none of them is:

**THEOREM 1.1.** *Let  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  be an ergodic flow. The following conditions are equivalent:*

- (i)  $\mathcal{T}$  is 2-DS (resp. 2-QS).
- (ii)  $T_t$  is 2-DS (resp. 2-QS) for  $t \in \mathbb{R}$  such that  $T_t$  is ergodic.
- (iii) There exists  $t_0 \in \mathbb{R}$  such that  $T_{t_0}$  is ergodic and  $T_{t_0}$  is 2-DS (resp. 2-QS).

Therefore, even though rather few explicit examples of 2-QS and 2-DS systems are known, these notions seem very natural. We first provide the proof of Theorem 1.1 for the 2-QS property. Then, building on the results obtained, we pass to the 2-DS property (see Sections 3.3.1 and 3.3.2). An essential role is played by lemmas which describe the relation between an extension of a flow and the corresponding extension of its time-one map.

---

<sup>(3)</sup> Another version of the R-property can be found in [79]; there however the notion was investigated from another point of view.

Roughly speaking, these extensions are of the same character (e.g. either both relatively distal or both relatively weakly mixing; for the definitions see Section 2.3). Our main tool is the so-called T-compactness property [47]. In the extended version of the paper (available on arXiv), we provide another, more direct proof of Theorem 1.1 in the case of the 2-DS property. It uses the notion of so-called *separating sieves* [63] which helps to avoid some technical difficulties.

**1.3.2. Mixing properties.** Recall the following classical result:

PROPOSITION 1.2. *Let  $\mathcal{T}$  be a measurable flow. Then  $\mathcal{T}$  is weakly mixing (resp. mildly mixing, lightly mixing, partially mixing, mixing, mixing of order  $n$ , rigid, partially rigid) if and only if  $T_1$  has the respective property.*

In the course of proving Theorem 1.1, as a by-product, we show that a “relative version” of the above proposition is true in the case of weak mixing. In fact, we show more. To state the result, we need to recall a classical theorem of Furstenberg and Zimmer (valid also for flows and other group actions):

THEOREM 1.3 (Furstenberg–Zimmer). *Let  $T$  be an ergodic automorphism on  $(X, \mathcal{B}, \mu)$  with a factor  $\mathcal{A} \subset \mathcal{B}$ . Then there exists a unique intermediate factor  $\mathcal{A} \subset \mathcal{C} \subset \mathcal{B}$  such that the extension  $\mathcal{B} \rightarrow \mathcal{C}$  is relatively weakly mixing and  $\mathcal{C} \rightarrow \mathcal{A}$  is relatively distal.*

The factorization from the above theorem is called *Furstenberg–Zimmer decomposition*. We show that for a flow and for its ergodic time- $t$  automorphism this decomposition is the same with respect to any flow factor <sup>(4)</sup>.

THEOREM 1.4. *Let  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}: (X, \mathcal{B}, \mu) \rightarrow (X, \mathcal{B}, \mu)$  be a flow with  $T_1$  ergodic and let  $\mathcal{A} \subset \mathcal{B}$  be  $\mathcal{T}$ -invariant. Then the Furstenberg–Zimmer decompositions for  $\mathcal{T}$  and for  $T_1$  with respect to  $\mathcal{A}$  are the same. In particular,  $\mathcal{T}$  is relatively weakly mixing (resp. relatively distal) over  $\mathcal{A}$  if and only if  $T_1$  is relatively weakly mixing (resp. relatively distal) over  $\mathcal{A}$ .*

**1.3.3. (Non-)uniqueness of the embedding.** Rudolph [68] gave an example of two non-isomorphic K-automorphisms with isomorphic squares (see also [67]). Also, Stepin and Eremenko [74] showed that for  $T$  with simple spectrum, which is embeddable into a measurable flow, either the embedding is unique, or  $T$  has infinitely many spectrally non-isomorphic embeddings. In a typical situation the set of possible embeddings (or roots of a given order) has cardinality continuum and no two different members are spectrally isomorphic.

In Section 4, we discuss the problem of uniqueness of the embedding from the point of view of self-joining properties. We show that for the time-one

---

<sup>(4)</sup> Again, an analogous result holds for automorphisms and their powers.

map of a flow with the MSJ property the embedding is unique. A similar phenomenon can be observed for time-one maps of flows having the R-property, in particular for horocycle flows. For flows with the MSJ property we provide two independent proofs. The first of them can be easily adapted to yield a similar result for roots. The second proof can be seen as a special case of a more general argument. Namely, we show that given a weakly mixing flow  $\mathcal{T}$ , a sufficient condition for the uniqueness of the embedding for  $T_1$  is that any ergodic joining of  $\mathcal{T}$  with a weakly mixing flow  $\mathcal{S}$  remains weakly mixing. This is also the core of the argument we apply to horocycle flows (using the theory of Ratner [64, 65, 66]).

Except for the Bernoulli shifts (both in the finite and the infinite entropy cases) [55, 56, 57, 59], ergodic automorphisms that are time- $t$  maps of MSJ flows or flows with the R-property <sup>(5)</sup> seem to be the first concrete examples of automorphisms with a unique embedding into a measurable flow.

Finally, we turn to “negative” examples. They are given in terms of compact group extensions of 2-fold simple systems. To construct a weakly mixing 2-fold simple automorphism which has two non-isomorphic square roots, we use the dihedral group of order 6 (one of these roots is 2-fold simple, the other is only 2-QS). To give an example of a weakly mixing automorphism which is 2-fold simple and embeddable into two non-isomorphic flows we use the group  $SU(2)$  (again, one of the flows obtained is 2-fold simple, the other is only 2-QS) <sup>(6)</sup>. Another natural class of automorphisms, which is of a different nature, consists of Gaussian automorphisms <sup>(7)</sup>.

**1.3.4. Counterexamples.** In Section 5, we focus on various types of counterexamples related to Theorem 1.1. The examples mentioned at the end of Subsection 1.3.3 show immediately that 2-fold simplicity of an (ergodic) time-one map of a flow does not imply that also the whole flow is 2-fold simple. Similarly, 2-fold simplicity of  $T^2$  does not imply 2-fold simplicity of  $T$ . In other words, Theorem 1.1 is “the most” we can hope for, and the aforementioned result of del Junco and Rudolph [20] cannot be reversed. In fact, already Danilenko [11] constructed an automorphism  $T$  such that  $T^2$  is 2-fold

<sup>(5)</sup> Recall that both flows with the MSJ and R-property are of zero entropy.

<sup>(6)</sup> Contrary to what was claimed in [50], it is possible to have non-uniqueness of embedding and roots also in the quasi-discrete spectrum case (cf. footnote 1). To see this, consider an irrational rotation on the circle  $Tx = x + \alpha$ . It has countably many non-isomorphic embeddings:  $T_{n,t}x = x + t(\alpha + n)$ . This also immediately yields countably many non-isomorphic roots for  $T$ . In a similar way, we can obtain non-isomorphic roots in the case of quasi-discrete, non-discrete spectrum. For  $T(x, y) = (x + \alpha, x + y)$  we have  $R \in C(T)$  for  $R(x, y) = (x, y + 1/2)$ , whence  $T^2 = (TR)^2$ .

<sup>(7)</sup> It was shown in [45] that JP systems are disjoint from Gaussian systems. Other related results can be found in [75, 18, 41]. We refer the reader e.g. to [10] or [53] for an introduction to Gaussian systems.

simple and  $T$  is only “2 : 1”. The key tool is the so-called (C, F)-construction procedure, which is an algebraic counterpart of the well-known cutting and stacking technique <sup>(8)</sup>. A general idea behind this type of construction is to use a larger (usually non-abelian) group with some special algebraic properties, which are reflected in the dynamics of a properly chosen subaction. The example given in [11] actually has even stronger properties than we need: the constructed automorphism is 2-QS (it is “2 : 1”), it is disjoint from simple systems, and its square is 2-fold simple. Notice that each of the actions constructed in Subsection 1.3.3 has a factor which is 2-fold simple. In particular, they are not disjoint from 2-fold simple systems. On the other hand, these constructions are much simpler than the example from [11].

The next example we consider has the same basic properties:  $T^2$  is 2-fold simple whereas  $T$  is only 2-QS. It is given in terms of a double group extension of an automorphism with the MSJ property (so it is again much simpler than the example from [11], but not disjoint from 2-fold simple automorphisms). The second group extension in our construction is very explicit: we use the affine cocycle. This is however not necessarily a strong point: it is unclear how to adjust this example to obtain a 2-QS flow with a 2-fold simple time-one map.

In Section 5.2, using the construction from [11] as the starting point, we provide an example of a flow  $\mathcal{T}$  with  $T_2$  ergodic and 2-fold simple, such that  $\mathcal{T}$  is only 2-QS (and not 2-fold simple). Moreover, the flow obtained has a factor with the same self-joining properties as the original flow and disjoint from 2-fold simple flows. To prove this, we show that given a 2-fold simple flow  $\mathcal{T}$  on  $(X, \mathcal{B}, \mu)$  with  $T_t$  ergodic and a flow factor  $\mathcal{A} \subset \mathcal{B}$ , either both  $\mathcal{T}|_{\mathcal{A}}$  and  $T_t|_{\mathcal{A}}$  are 2-fold simple or none of them is. In particular, the horocycle flows and their factors cannot serve as an example here. In our (C, F)-construction we deal with uncountable non-abelian groups. This results in additional technical difficulties related to equidistribution of sequences (cf. [12, 13]); since this part of the paper is of a different flavor than the rest of it, we include the necessary tools in Appendix A.

**2. Preliminaries.** We deal with measure-preserving automorphisms and flows, i.e.  $\mathbb{Z}$ - and  $\mathbb{R}$ -actions on standard Borel spaces, preserving a probability measure. In particular, we assume that all flows are measurable, i.e. the map  $X \times \mathbb{R} \ni (x, t) \mapsto T_t x \in X$  is measurable. Sometimes the invariant measure is not fixed and we work with a standard Borel space  $(X, \mathcal{B})$  (the existence of an invariant measure will always be guaranteed). The properties which are of our interest are invariant under measure-theoretical isomorphisms.

---

<sup>(8)</sup> This method is a useful tool for producing examples and counterexamples of different types of behavior. For a detailed survey, see [12].

Therefore, without loss of generality, we can make a tacit assumption that the flows we consider are continuous flows on compact metric spaces <sup>(9)</sup>.

We formulate most of the definitions in terms of  $\mathbb{Z}$ -actions. However, they can be transferred directly (or almost directly) to other group actions.

**2.1. Topology, invariant measures, ergodic decomposition.** Let  $X$  be a compact metric space. Denote by  $M_1(X)$  the set of probability measures on  $X$ . With the usual weak topology,  $M_1(X)$  is compact and metrizable. For a flow  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  and automorphism  $T$  on  $X$  we define

$$\begin{aligned} \mathcal{P}_{\mathbb{R}}(\mathcal{T}) &:= \{\mu \in M_1(X) : \mu \circ T_t = \mu \text{ for all } t \in \mathbb{R}\}, \\ \mathcal{P}_{\mathbb{R}}^e(\mathcal{T}) &:= \{\mu \in \mathcal{P}_{\mathbb{R}}(\mathcal{T}) : \mu \text{ is ergodic for } \mathcal{T}\}, \\ \mathcal{P}_{\mathbb{Z}}(T) &:= \{\mu \in M_1(X) : \mu \circ T = \mu\}, \\ \mathcal{P}_{\mathbb{Z}}^e(T) &:= \{\mu \in \mathcal{P}_{\mathbb{Z}}(T) : \mu \text{ is ergodic for } T\}. \end{aligned}$$

REMARK 2.1. Since  $X$  is a compact metric space and  $\mathcal{T}$  is assumed to be continuous, it follows that the map  $\mathbb{R} \ni t \mapsto \mu \circ T_t \in M_1(X)$  is continuous whenever  $\mathcal{P}_{\mathbb{R}}(\mathcal{T}) \neq \emptyset$ .

We will identify the following spaces:  $\mathbb{R}/\mathbb{Z}$ ,  $[0, 1)$ ,  $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$  ( $[0, 1)$  will be understood additively and  $\mathbb{T}$  multiplicatively), and we will equip them with the Lebesgue measure. Furthermore, given a measurable flow  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  and a measure  $\nu \in \mathcal{P}_{\mathbb{Z}}^e(T_1)$  we set

$$(2.1) \quad \underline{\nu} = \int_0^1 \nu \circ T_t dt.$$

Recall the following folklore result (see, e.g., [11]):

LEMMA 2.2. *The measure  $\underline{\nu}$  is  $\mathcal{T}$ -invariant and ergodic. Moreover, the map*

$$(2.2) \quad \mathcal{P}_{\mathbb{Z}}^e(T_1) \ni \nu \mapsto \underline{\nu} \in \mathcal{P}_{\mathbb{R}}^e(\mathcal{T})$$

*is onto and (2.1) is the ergodic decomposition of  $\underline{\nu}$  for  $T_1$ .*

Note that the measures  $\nu \circ T_t$  in (2.1) need not be pairwise orthogonal. However, we have the following:

LEMMA 2.3. *In the decomposition (2.1) either all measures  $\nu \circ T_t$  are equal to  $\nu$  or there exists  $k \geq 1$  such that*

$$(2.3) \quad \underline{\nu} = \int_{\mathbb{T}} \nu \circ T_t dt = k \int_0^{1/k} \nu \circ T_t dt$$

*and for  $0 \leq t < 1/k$  all measures  $\nu \circ T_t$  are mutually singular.*

---

<sup>(9)</sup> In the case of flows one uses the special flow representation [2, 3] and a result from [38] on the homology class of integrable functions.



*Proof.* Notice that any two measures in (2.1) are either equal or mutually singular. Let  $A := \{t \in \mathbb{T} : \eta = \eta \circ T_t\}$ . This is clearly a subgroup of  $\mathbb{T}$ . Moreover,  $A$  is closed by Remark 2.1. Therefore either  $A = \mathbb{T}$  or  $A$  is finite. In the latter case we obtain formula (2.3) with  $k = \#A$ . ■

REMARK 2.4. If  $\mathcal{T}$  is a suspension flow with the first return time to the base equal to  $a > 0$  and a weakly mixing first return time map, then the discrete part of the spectrum is concentrated on  $(1/a)\mathbb{Z}$ .

REMARK 2.5. Recall that a flow is a suspension flow (with a constant first return time to the base) if and only if it has a rational point in its discrete spectrum [82]. Therefore, whenever (2.3) holds, the flow  $\mathcal{T}$  can be represented as a suspension flow over its  $1/k$ -time map with common return time to the base equal to  $1/k$ . In other words, the whole space can be identified with  $(X \times [0, 1/k], \nu \otimes \lambda_{[0, 1/k]})$ , where we glue together each pair of points  $(x, 1/k)$  and  $(T_{1/k}x, 0)$ , with the action of  $(T_t)_{t \in \mathbb{R}}$  given by  $T_t(x, s) = (x, s + t)$ . Moreover, by Remark 2.4, the discrete part of the spectrum is concentrated on  $k\mathbb{Z}$ .

**2.2. Joinings I: Basic definitions and properties.** Let  $T$  and  $S$  be automorphisms of  $(X, \mathcal{B}, \mu)$  and  $(Y, \mathcal{C}, \nu)$  respectively. By  $\mathcal{J}(T, S)$  we denote the set of all *joinings* between  $T$  and  $S$ , i.e. the set of all  $T \times S$ -invariant probability measures on  $(X \times Y, \mathcal{B} \otimes \mathcal{C})$  whose projections onto  $X$  and  $Y$  are equal to  $\mu$  and  $\nu$  respectively. For  $\mathcal{J}(T, T)$  we write  $\mathcal{J}(T)$  and we speak about *self-joinings*. If we assume additionally that both  $T$  and  $S$  are ergodic, then  $\mathcal{J}(T, S)$  is a simplex whose extreme points are joinings ergodic for  $T \times S$ ; we then write  $\mathcal{J}^e(T, S)$  (or  $\mathcal{J}^e(T)$  if  $S = T$ ). If  $\mathcal{J}(T, S) = \{\mu \otimes \nu\}$  we say that  $T$  and  $S$  are *disjoint* [26] and write  $T \perp S$ .

In a similar way one defines joinings of higher order, i.e. joinings of more than two automorphisms. We denote them by  $\mathcal{J}(T_1, \dots, T_k)$ ,  $k \geq 2$ , and  $\mathcal{J}(T_1, T_2, \dots)$  in the case of a finite and an infinite number of automorphisms respectively. Again, we may be interested in the situation where all the  $T_i$  are isomorphic and speak about self-joinings. To distinguish self-joinings of different orders we write e.g.  $\mathcal{J}_\infty^e(T)$  for infinite ergodic self-joinings of  $T$  or  $\mathcal{J}_3(T)$  for three-fold self-joinings of  $T$ . By  $\mathcal{J}_2(T; \mathcal{A})$ , where  $\mathcal{A}$  is a factor of  $T$ , we denote all 2-fold self-joinings of  $T$  which project onto  $\mathcal{A} \otimes \mathcal{A}$  as the diagonal measure. Sometimes, to simplify the notation, we will denote a joining between  $T$  and  $S$  by  $T \vee S$ .

Recall that 2-fold joinings are in one-to-one correspondence with Markov operators  $\Phi: L^2(X, \mathcal{B}, \mu) \rightarrow L^2(Y, \mathcal{C}, \nu)$  such that  $\Phi \circ T = S \circ \Phi$ . This identification allows us to view  $\mathcal{J}_2(T)$  as a metrizable compact semitopological semigroup endowed with the weak operator topology. A metric compatible with the weak topology on  $\mathcal{J}_2(T)$  can be defined in the following way. Fix a dense subset  $\{f'_i : i \geq 1\} \subset L^2(X, \mathcal{B}, \mu)$ , set  $f_i := f'_i / \|f'_i\|_{L^2(X, \mu)}$  for  $i \geq 1$

and let  $\bar{d}_{\{f_n\}_{n \in \mathbb{N}}}$  be given by

$$(2.4) \quad \bar{d}_{\{f_n\}_{n \in \mathbb{N}}}(\xi, \nu) = \sum_{i,j=1}^{\infty} \frac{1}{2^{i+j}} \left| \int_{X \times X} f_i(x) \overline{f_j(y)} d\xi(x, y) - \int_{X \times X} f_i(x) \overline{f_j(y)} d\nu(x, y) \right|.$$

Notice that for  $\xi, \nu \in \mathcal{J}_2(T)$  we always have

$$(2.5) \quad \bar{d}_{\{f_n\}_{n \in \mathbb{N}}}(\xi \circ (T \times T), \nu \circ (T \times T)) = \bar{d}_{(f_n)_{n \in \mathbb{N}}}(\xi, \nu).$$

REMARK 2.6. We will often use Lemmas 2.2 and 2.3 for joinings. Notice that given a flow  $\mathcal{T}$  with  $T_1$  ergodic, the map  $\eta \mapsto \underline{\eta}$  has values in  $\mathcal{J}_2^e(\mathcal{T})$  whenever  $\eta \in \mathcal{J}_2^e(T_1)$ . Moreover, it is surjective.

Recall that  $S$  is a *factor* of  $T$  (or  $T$  is an *extension* of  $S$ ) if there exists  $\pi: X \rightarrow Y$  such that  $\pi \circ T = S \circ \pi$  and  $\nu = \pi_*(\mu)$ , i.e. for any  $A \in \mathcal{C}$  we have  $\mu(\pi^{-1}A) = \nu(A)$ . We then write  $T \rightarrow S$ . Notice that if  $S$  is a factor of  $T$  then  $\pi^{-1}(\mathcal{C}) \subset \mathcal{B}$  is  $T$ -invariant. On the other hand, every  $T$ -invariant sub- $\sigma$ -algebra  $\mathcal{A} \subset \mathcal{B}$  corresponds to some factor of  $T$ . If no confusion arises we will write  $\mathcal{B} \rightarrow \mathcal{A}$ .

We will use the following types of joinings:

- For  $S_2, \dots, S_N \in C(T)$  <sup>(10)</sup> the measure  $\nu = \mu_{S_2, \dots, S_N}$  defined by

$$\mu_{S_2, \dots, S_N}(A_1 \times \dots \times A_N) = \mu(A_1 \cap S_2^{-1}A_2 \cap \dots \cap S_N^{-1}(A_N))$$

(with an obvious modification for  $N = \infty$ ) is an  $N$ -self-joining of  $T$  called a *graph self-joining* (or *off-diagonal self-joining*). If  $S_2 = \dots = S_N = \text{Id}$ , we speak about the *diagonal self-joining* and we denote it by  $\Delta_N$ . For  $N = 2$  we often write  $\Delta$  instead of  $\Delta_2$ . When confusion may arise, we denote the marginals of  $\Delta$  in parentheses:  $\Delta(\mu)$ .

- Each pair of factors  $\mathcal{A}_1, \mathcal{A}_2 \subset \mathcal{B}$ , together with  $\lambda \in \mathcal{J}(T|_{\mathcal{A}_1}, T|_{\mathcal{A}_2})$ , yields a self-joining of  $T$  defined by

$$\widehat{\lambda}(A_1 \times A_2) = \int_{X/\mathcal{A}_1 \times X/\mathcal{A}_2} E(A_1|\mathcal{A}_1)(\bar{x})E(A_2|\mathcal{A}_2)(\bar{y}) d\lambda(\bar{x}, \bar{y}).$$

It is called a *relatively independent extension* of  $\lambda$ . In particular, when  $\mathcal{A}_1 = \mathcal{A}_2 = \mathcal{A}$  and  $\lambda = \Delta$  we say that  $\widehat{\Delta}$  is a relatively independent extension *over*  $\mathcal{A}$ . We denote it by  $\mu \otimes_{\mathcal{A}} \mu$ .

- Let  $\lambda \in \mathcal{J}(T_1, \dots, T_n)$ . Then  $\lambda|_{X_i, X_j} = \lambda|_{X_i \times X_j}$  defined by

$$\lambda|_{X_i \times X_j}(C \times D) = \lambda(\underbrace{X \times \dots \times X}_{i-1} \times C \times \underbrace{X \times \dots \times X}_{j-i-1} \times D \times \underbrace{X \times \dots \times X}_{n-j})$$

is an element of  $\mathcal{J}(T_i, T_j)$ . Moreover, if  $\lambda \in \mathcal{J}^e(T_1, \dots, T_n)$  then  $\lambda|_{X_i, X_j} \in \mathcal{J}^e(T_i, T_j)$ .

<sup>(10)</sup>  $C(T)$  stands for the *centralizer* of  $T$ , i.e. the group of all measure-preserving automorphisms which commute with  $T$ .

• Let  $\lambda \in \mathcal{J}(T_1, T_2)$  and let  $\mathcal{A}_1, \mathcal{A}_2$  be factors of  $T_1, T_2$ , respectively. Then  $\lambda|_{\mathcal{A}_1 \otimes \mathcal{A}_2} \in \mathcal{J}((T_1)|_{\mathcal{A}_1}, (T_2)|_{\mathcal{A}_2})$ .

**2.3. Factors and extensions.** Any ergodic extension  $\tilde{T} \rightarrow T$  is isomorphic to a *skew product* over  $T$ , i.e. to an automorphism  $\bar{T}$  on  $(X \times Y, \mathcal{B} \otimes \mathcal{C}, \mu \otimes \nu)$  of the form

$$(2.6) \quad \bar{T}(x, y) = (Tx, S_x y),$$

where  $(S_x)_{x \in X}$  is a measurable family of automorphisms of some standard probability space  $(Y, \mathcal{C}, \nu)$  (see [1]). A particular case of skew products are *Rokhlin extensions*, i.e. automorphisms of the form  $T_{S, \varphi}(x, y) = (Tx, S_{\varphi(x)}y)$ , where  $(S_g)_{g \in G}$  is a measurable representation of a locally compact abelian group  $G$  in the group of automorphisms of  $(Y, \mathcal{C}, \nu)$ , and  $\varphi: X \rightarrow G$  is a measurable function.

**2.3.1. Relatively weakly mixing extensions.** Following [27], we say that  $T: (X, \mathcal{B}, \mu) \rightarrow (X, \mathcal{B}, \mu)$  is *relatively weakly mixing* with respect to factor  $\mathcal{A} \subset \mathcal{B}$  if the relatively independent extension of the diagonal self-joining of  $T|_{\mathcal{A}}$  is ergodic. In terms of the skew product representation (2.6), the property of relative weak mixing is equivalent to the ergodicity of the automorphism  $\bar{\bar{T}}$  acting on  $(X \times Y \times Y, \mathcal{B} \otimes \mathcal{C} \otimes \mathcal{C}, \mu \otimes \nu \otimes \nu)$ , given by  $\bar{\bar{T}}(x, y_1, y_2) = (Tx, S_x y_1, S_x y_2)$ .

**2.3.2. Relatively isometric and relatively distal extensions.** Relative distality is a concept of “opposite nature” to relative weak mixing (cf. Theorem 1.3).

*Compact group extensions.* Let  $G$  be a compact metrizable group with the normalized Haar measure  $\lambda_G$  and let  $\varphi: X \rightarrow G$  be measurable. Then  $T_\varphi: (x, g) \mapsto (Tx, \varphi(x)g)$  is an automorphism of  $(X \times G, \mathcal{B} \otimes \mathcal{B}(G), \mu \otimes \lambda_G)$ . It is called a (compact) *group extension* of  $T$ . We have the following:

PROPOSITION 2.7 (see e.g. [20]). *If  $T: (X, \mathcal{B}, \mu) \rightarrow (X, \mathcal{B}, \mu)$  is an ergodic compact group extension of its factor  $\mathcal{A}$  then  $J_2^e(T; \mathcal{A})$  consists of graph joinings.*

PROPOSITION 2.8 ( $\mathbb{Z}$ -actions: [75, 78]; actions of other groups: [20]). *If  $T: (X, \mathcal{B}, \mu) \rightarrow (X, \mathcal{B}, \mu)$  is ergodic,  $\mathcal{A}$  is a factor of  $T$  and  $J_2^e(T; \mathcal{A})$  consists of graph joinings then  $\mathcal{B} \rightarrow \mathcal{A}$  is a compact group extension.*

LEMMA 2.9 ([43]). *Let  $T: (X, \mathcal{B}, \mu) \rightarrow (X, \mathcal{B}, \mu)$  be a (not necessarily ergodic) compact group extension of an ergodic factor  $\mathcal{A}$ . Then every ergodic component of  $T$  is also isomorphic to some compact group extension of  $\mathcal{A}$ .*

*Isometric extensions.* Let  $G$  be a compact group with a closed subgroup  $H$  and let  $\varphi: X \rightarrow G$  be measurable. The action of  $T_\varphi$  restricted to  $X \times G/H$  is called an *isometric extension* of  $T$ .

In [81] Zimmer introduced the notion of relatively discrete spectrum. It is based on the classical notion of discrete spectrum of a unitary representation. As we will not need the precise definition, let us only remark that

(2.7) an ergodic extension has relatively discrete spectrum if and only if it is isometric.

*Distal extensions.* We say that an extension  $\tilde{T}: (\tilde{X}, \tilde{\mathcal{B}}, \tilde{\mu}) \rightarrow (X, \mathcal{B}, \mu)$  is *relatively distal* [80] if there exists a transfinite sequence  $(\mathcal{B}_\alpha)_{\alpha \leq \beta}$  of factors of  $\tilde{T}$  such that  $\mathcal{B}_0 = \mathcal{B}$ ,  $\mathcal{B} = \tilde{\mathcal{B}}$ , the extensions  $\mathcal{B}_{\alpha+1} \rightarrow \mathcal{B}_\alpha$  are isometric and, for  $\alpha$  being a limit ordinal,  $\mathcal{B}_\alpha$  is the inverse limit of the preceding factors.

**2.4. Joinings II: Self-joining properties.** Let  $T: (X, \mathcal{B}, \mu) \rightarrow (X, \mathcal{B}, \mu)$  be an ergodic automorphism. We say that:

- $T$  has the *pairwise independence property* (PID) if every pairwise independent self-joining of  $T$  is equal to the product measure;
- $T$  is *2-fold simple* if  $\mathcal{J}_2^e(T) \subset \{\mu_S : S \in C(T)\} \cup \{\mu \otimes \mu\}$ ;  $T$  is *simple* if it is 2-fold simple and PID;
- $T$  has *minimal self-joinings* (MSJ) if it is 2-fold simple and  $C(T) = \{T^k : k \in \mathbb{Z}\}$ ;
- $T$  is *2-fold quasi-simple* (2-QS) if  $\mathcal{J}^e(T) \setminus \{\mu \otimes \mu\}$  consists of isometric extensions of each of the coordinate factors;  $T$  is *quasi-simple* if it is 2-QS and PID;
- $T$  is “ $n : 1$ ” for some  $n \in \mathbb{N}$  if  $\mathcal{J}^e(T) \setminus \{\mu \otimes \mu\}$  consists of “ $n : 1$ ”-extensions of each of the coordinate factors (it is a special case of 2-QS);
- $T$  is *2-fold distally simple* (2-DS) if  $\mathcal{J}^e(T) \setminus \{\mu \otimes \mu\}$  consists of distal extensions of each of the coordinate factors;  $T$  is *distally simple* if it is 2-DS and PID.

We say that  $T$  has the *joining primeness property* (JP) if for every pair of weakly mixing systems  $S_1, S_2$  and every  $\lambda \in \mathcal{J}^e(T, S_1 \times S_2)$  we have  $\lambda = \lambda|_{X, Y_1} \otimes \nu_2$  or  $\lambda = \lambda|_{X, Y_2} \otimes \nu_1$ . Notice that assuming  $S_1$  and  $S_2$  are isomorphic would give an equivalent definition. We will show later that JP is, in fact, an intrinsic notion and we will provide an equivalent definition in terms of self-joinings.

Finally, we mention *Ratner’s property* (R-property) of flows. It has a different flavor than all the properties mentioned above. Since the definition is very technical and will not be used explicitly, we will not quote it here. Instead, we refer the reader to [25, 66, 10, 24] for the details. Note that although the definition of R-property does not refer to joinings, it imposes

restrictions on the joinings of the flow under consideration. The following proposition is a little imprecise, but it will be sufficient for our needs.

PROPOSITION 2.10 ([25]). *Let  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  be a weakly mixing flow on  $(X, \mathcal{B}, \mu)$  which has the  $R$ -property. Let  $\mathcal{S} = (S_t)_{t \in \mathbb{R}}$  be an ergodic flow on  $(Y, \mathcal{C}, \nu)$  and let  $\lambda \in \mathcal{J}^e(\mathcal{T}, \mathcal{S})$ . Then either  $\lambda = \mu \otimes \nu$  or  $\lambda$  is a finite extension of  $\nu$ .*

**2.5. Distal and isometric extensions:  $\mathbf{T}$ -compactness.** Let  $\mathcal{A}$  be a factor of  $T$ . Following [47], we say that an extension  $\mathcal{B} \rightarrow \mathcal{A}$  is  $T$ -compact if for any  $\varepsilon > 0$  there exists  $N \geq 2$  such that for each  $\nu \in \mathcal{J}_N^e(T; \mathcal{A})$  we can find  $1 \leq i < j \leq N$  such that  $\bar{d}(\nu|_{X_i \times X_j}, \Delta_N|_{X_i \times X_j}) < \varepsilon$  <sup>(11)</sup>.

PROPOSITION 2.11 ([47]). *Assume that  $T$  is an ergodic automorphism on  $(X, \mathcal{B}, \mu)$  and let  $\mathcal{A} \subset \mathcal{B}$  be its factor. Then the following are equivalent:*

- (i)  $\mathcal{B} \rightarrow \mathcal{A}$  has relatively discrete spectrum,
- (ii)  $\mathcal{B} \rightarrow \mathcal{A}$  is isometric,
- (iii)  $\mathcal{B} \rightarrow \mathcal{A}$  is  $T$ -compact.

### 3. Self-joining properties of $\mathcal{T}$ and $T_1$ , part I

**3.1. Simplicity.** Our starting point is a result of del Junco and Rudolph. We rephrase it slightly and also include a proof, using the language compatible with that used later on.

PROPOSITION 3.1 ([20]). *Let  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  be a weakly mixing flow. The following are equivalent:*

- (i)  $\mathcal{T}$  is 2-fold simple,
- (ii)  $T_1$  is 2-fold simple and  $C(\mathcal{T}) = C(T_1)$ .

REMARKS. (1) In particular,  $C(\mathcal{T}) = C(T_1)$  whenever  $C(T_1)$  is abelian.

(2) An analogous result to Proposition 3.1 holds for automorphisms and their powers.

(3) A flow  $\mathcal{T}$  which is not weakly mixing, and which has an ergodic time  $T_t$  which is 2-fold simple, is also 2-fold simple. Indeed, as  $\mathcal{T}$  is not weakly mixing,  $T_t$  is not weakly mixing either, whence it has purely discrete spectrum [20]. Therefore  $\mathcal{T}$  also has purely discrete spectrum, and in particular, it is 2-fold simple.

*Proof of Proposition 3.1.* Suppose that  $\mathcal{T}$  is 2-fold simple. It suffices to show that  $\mathcal{J}_2^e(\mathcal{T}) = \mathcal{J}_2^e(T_1)$ . We have  $\mathcal{J}_2^e(\mathcal{T}) \subset \mathcal{J}_2^e(T_1)$ . Suppose that we can

---

<sup>(11)</sup> See the definitions in Section 2.2.

find  $\eta \in \mathcal{J}_2^e(T_1) \setminus \mathcal{J}_2^e(\mathcal{T})$ . Define

$$(3.1) \quad \underline{\eta} = \int_0^1 \eta \circ (T_t \times T_t) dt.$$

By Remark 2.6, we have  $\underline{\eta} \in \mathcal{J}_2^e(\mathcal{T}) \subset \mathcal{J}_2^e(T_1)$ . It follows by Lemma 2.3 and by the uniqueness of ergodic decomposition that (3.1) takes the form  $\underline{\eta} = \eta$ , which contradicts the choice of  $\eta$ .

The other implication follows easily: we obtain  $\mathcal{J}_2^e(T_1) \subset \mathcal{J}_2^e(\mathcal{T})$ , whence  $\mathcal{J}_2(T_1) \subset \mathcal{J}_2(\mathcal{T})$ . The reverse inclusion is obvious. ■

REMARK 3.2. Let  $T$  and  $S$  be 2-fold simple such that  $T^k = S^l$  for some  $k, l \in \mathbb{Z} \setminus \{0\}$ . It follows from Proposition 3.1 that  $T^k = S^l$  is 2-fold simple and  $C(T) = C(T^k) = C(S^l) = C(S)$ . In particular,  $\mathcal{J}_2^e(T) = \mathcal{J}_2^e(S)$ .

For a measure-preserving flow  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  let

$$I_s(\mathcal{T}) = \{t \in \mathbb{R} : T_t \text{ is 2-fold simple}\} \cup \{0\}.$$

LEMMA 3.3. *Let  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  be a weakly mixing flow. Suppose that  $T_\alpha$  and  $T_\beta$  are 2-fold simple for some  $\alpha, \beta \neq 0$  such that  $\alpha/\beta \notin \mathbb{Q}$ . Then  $\mathcal{T}$  is 2-fold simple. In particular,  $I_s(\mathcal{T}) = \mathbb{R}$ .*

*Proof.* We may assume without loss of generality that  $\beta = 1$  and  $\alpha \notin \mathbb{Q}$ . There are two possibilities: (i)  $C(T_1) = C(T_\alpha)$ , (ii)  $C(T_1) \neq C(T_\alpha)$ .

In case (i), fix  $t_0 \in \mathbb{R}$ . For  $n \in \mathbb{N}$  choose  $t_n \in \mathbb{Z} + \alpha\mathbb{Z}$  such that  $t_n \rightarrow t_0$  (this can always be done as  $\alpha \notin \mathbb{Q}$ ). Take  $S \in C(T_1) = C(T_\alpha)$ . We have

$$ST_{t_n} \rightarrow ST_{t_0} \quad \text{and} \quad T_{t_n}S \rightarrow T_{t_0}S.$$

Since  $ST_{t_n} = T_{t_n}S$ , it follows that  $ST_{t_0} = T_{t_0}S$ , i.e.  $C(T_1) = C(\mathcal{T})$ . By Proposition 3.1,  $\mathcal{T}$  is 2-fold simple.

In case (ii), we may assume that there exists  $S \in C(T_1) \setminus C(T_\alpha)$ . Consider

$$(3.2) \quad \underline{\mu_S} := \int_0^1 \mu_S \circ (T_t \times T_t) dt.$$

By Remark 2.6,  $\underline{\mu_S} \in \mathcal{J}_2^e(\mathcal{T})$ . Clearly  $\underline{\mu_S} \notin \mathcal{J}_2^e(T_1)$  (otherwise the measures in the decomposition (3.2) would be all equal, which would imply that  $S \in C(\mathcal{T}) \subset C(T_\alpha)$ ). By Lemma 2.3, the ergodic decomposition of  $\underline{\mu_S}$  for  $T_1 \times T_1$  is

$$(3.3) \quad \underline{\mu_S} = k \int_0^{1/k} \mu_S \circ (T_t \times T_t) dt$$

for some  $k \geq 1$ . By Remark 2.5,

$$(3.4) \quad \text{the set of eigenvalues of } (\mathcal{T} \times \mathcal{T}, \underline{\mu_S}) \text{ is equal to } k\mathbb{Z}.$$

Suppose that  $\underline{\mu}_{\mathcal{S}}$  is not ergodic for  $T_\alpha \times T_\alpha$ . Then, by Lemma 2.2,  $\underline{\mu}_{\mathcal{S}} = (1/\alpha) \int_0^\alpha \mu_W \circ (T_t \times T_t) dt$  for some  $W \in C(T_\alpha)$ . By Lemma 2.3 the ergodic decomposition of  $\underline{\mu}_{\mathcal{S}}$  for  $T_\alpha \times T_\alpha$  is therefore of the form  $\underline{\mu}_{\mathcal{S}} = (l/\alpha) \int_0^{\alpha/l} \mu_W \circ (T_t \times T_t) dt$  for some  $l \geq 1$ . Hence, by Remark 2.5, the set of eigenvalues  $(\mathcal{T} \times \mathcal{T}, \underline{\mu}_{\mathcal{S}})$  is equal to  $(l/\alpha)\mathbb{Z}$ . This contradicts (3.4).

Suppose now that  $\underline{\mu}_{\mathcal{S}}$  is ergodic for  $T_\alpha \times T_\alpha$ . Then  $\underline{\mu}_{\mathcal{S}} = \mu_W$  for some  $W \in C(T_\alpha)$ . It follows that  $(T_\alpha \times T_\alpha, \underline{\mu}_{\mathcal{S}})$  is weakly mixing, whence also  $(\mathcal{T} \times \mathcal{T}, \underline{\mu}_{\mathcal{S}})$  is weakly mixing. This contradicts (3.4) again. ■

**PROPOSITION 3.4.** *Let  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  be a weakly mixing flow. Then  $I_s(\mathcal{T}) = \mathbb{R}$  or  $I_s(\mathcal{T}) = \beta\mathbb{Z}$  for some  $\beta \in \mathbb{R}$ . Moreover, if  $I_s(\mathcal{T}) = \mathbb{R}$  then  $\mathcal{T}$  is 2-fold simple.*

*Proof.* Suppose that  $I_s(\mathcal{T}) \neq \beta\mathbb{Z}$  for all  $\beta \in \mathbb{R}$ . Then either (i) there exist  $t_1, t_2 \in I_s(\mathcal{T}) \setminus \{0\}$  which are rationally independent, or (ii) all numbers in  $I_s(\mathcal{T})$  are rationally dependent.

In case (i) it follows from Lemma 3.3 that  $\mathcal{T}$  is 2-fold simple and  $I_s(\mathbb{R}) = \mathbb{R}$ . In case (ii) there exists  $\alpha > 0$  such that  $I_s(\mathcal{T}) \subset \alpha\mathbb{Q}$ . Take  $t_1, t_2 \in I_s(\mathcal{T})$ . We claim that the whole additive subgroup generated by  $t_1$  and  $t_2$  is in  $I_s(\mathcal{T})$ . Indeed,  $t_1 = p_1/q$ ,  $t_2 = p_2/q$  for some  $p_1, p_2, q \in \mathbb{Z}$ . Let  $d = \gcd(p_1, p_2)$ . Then for some  $k, l \in \mathbb{Z}$  we have  $kp_1 + lp_2 = d$ . Hence  $d/q = kp_1/q + lp_2/q$ . It follows by Remark 3.2 that  $C(T_{p_1/q}) = C(T_{p_2/q})$ . Therefore  $C(T_{p_1/q}) \subset C(T_{d/q})$ . On the other hand,  $C(T_{d/q}) \subset C(T_{p_1/q})$  since  $T_{d/q}$  is a root of  $T_{p_1/q}$ . Hence  $C(T_{d/q}) = C(T_{p_1/q})$  and it follows that  $T_{d/q}$  is 2-fold simple. Therefore, for all  $k \in \mathbb{Z}$  also  $T_{kd/q}$  is 2-fold simple. It follows that  $I_s(\mathcal{T}) = I_s(\mathcal{T}) \cap \alpha\mathbb{Q}$  is an additive subgroup of  $\alpha\mathbb{Q}$ . Notice that  $I_s(\mathcal{T})$  cannot be dense in  $\mathbb{R}$ . Indeed, if it were, an argument similar to the one in case (i) in the proof of Lemma 3.3 would imply  $I_s(\mathcal{T}) = \mathbb{R}$ . ■

**3.2. Joining primeness property.** We will now show that the converse of Proposition 3.1 is true if we replace “2-fold simple” with “JP”:

**PROPOSITION 3.5.** *Let  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  be a measure-preserving flow with  $T_1$  ergodic. If  $T_1$  has the JP property then  $\mathcal{T}$  has the JP property.*

*Proof.* Let  $\underline{\eta} \in J^e(\mathcal{T}, \mathcal{R} \times \mathcal{S}) \subset J(T_1, R_1 \times S_1)$  for some weakly mixing flows  $\mathcal{R}$  and  $\mathcal{S}$ . If  $\underline{\eta} \in J^e(T_1, R_1 \times S_1)$  then  $\underline{\eta} = \underline{\eta}|_{X \times Y} \otimes \nu_2$  or  $\underline{\eta} = \underline{\eta}|_{X \times Z} \otimes \nu_1$ . If  $\underline{\eta} \notin J^e(T_1, R_1 \times S_1)$  then by Lemmas 2.2, 2.3 and Remark 2.6 we may assume that for some  $\eta \in J^e(T_1, R_1 \times S_1)$  we have

$$\underline{\eta} = \int_0^1 \eta \circ (T_t \times R_t \times S_t) dt.$$

Notice that  $\eta$  cannot be the product measure. Indeed, it would then be invariant under the product flow  $\mathcal{R} \times \mathcal{S} = (R_t \times S_t)_{t \in \mathbb{R}}$  and equal to  $\underline{\eta}$ , which

is impossible since  $\underline{\eta} \notin \mathcal{J}^e(T_1, R_1 \times S_1)$ . Therefore, up to a permutation of the coordinates, we have  $\eta = \eta|_{X \times Y} \otimes \nu_2$ . Hence

$$\underline{\eta} = \int_0^1 \eta|_{X \times Y} \circ (T_t \times R_t) \otimes \nu_2 dt = \left( \int_0^1 \eta|_{X \times Y} \circ (T_t \times R_t) dt \right) \otimes \nu_2. \blacksquare$$

REMARK 3.6. If  $\mathcal{T}^{(1)}$  and  $\mathcal{T}^{(2)}$  are weakly mixing and all ergodic joinings of  $\mathcal{T}^{(1)}$  and  $\mathcal{T}^{(2)}$  are weakly mixing then  $\mathcal{J}(\mathcal{T}^{(1)}, \mathcal{T}^{(2)}) = \mathcal{J}(T_1^{(1)}, T_1^{(2)})$  and  $\mathcal{J}^e(\mathcal{T}^{(1)}, \mathcal{T}^{(2)}) = \mathcal{J}^e(T_1^{(1)}, T_1^{(2)})$ . Indeed, by Lemma 2.2 the only possibility to obtain a new ergodic joining is via suspension, but every suspension has a discrete spectrum factor.

REMARK 3.7. If  $\mathcal{J}_2(\mathcal{T}) = \mathcal{J}_2(T_1)$  then  $\mathcal{T}$  and  $T_1$  have the same invariant  $\sigma$ -algebras and  $C(\mathcal{T}) = C(T_1)$ .

Given an ergodic automorphism  $R$  on a standard probability Borel space  $(Z, \mathcal{D}, \rho)$ , we denote by  $\mathcal{JF}(R)$  the family of all factors of all infinite ergodic self-joinings of  $R$ . Notice that

(3.5)  $\mathcal{JF}(R)$  is closed under (infinite) ergodic self-joinings and factors.

Therefore, for each ergodic automorphism  $S$  acting on  $(Y, \mathcal{C}, \nu)$  there exists a largest factor  $\mathcal{A}_R(S) \subset \mathcal{C}$  of  $S$  such that  $\mathcal{A}_R(S) \in \mathcal{JF}(R)$  <sup>(12)</sup>. Since  $S|_{W_{\mathcal{A}_R(S)}}$  and  $S|_{\mathcal{A}_R(S)}$  are isomorphic for each  $W \in C(S)$ , we find that

(3.6)  $\mathcal{A}_R(S)$  is invariant under all  $W \in C(S)$ .

Now, fix a weakly mixing automorphism  $R: (Z, \mathcal{D}, \rho) \rightarrow (Z, \mathcal{D}, \rho)$  and an ergodic automorphism  $S: (Y, \mathcal{C}, \nu) \rightarrow (Y, \mathcal{C}, \nu)$ . Then, by (3.6),  $\mathcal{A}_{R \times R}(S \times S)$  is a factor of the product action  $(S^i \times S^j)_{(i,j) \in \mathbb{Z}^2}$ . Recall a result on product  $\mathbb{Z}^2$ -actions:

PROPOSITION 3.8 ([6]). *Let  $\tilde{T}: (\tilde{X}, \tilde{\mathcal{B}}, \tilde{\mu}) \rightarrow (\tilde{X}, \tilde{\mathcal{B}}, \tilde{\mu})$  and  $\tilde{S}: (\tilde{Y}, \tilde{\mathcal{C}}, \tilde{\nu}) \rightarrow (\tilde{Y}, \tilde{\mathcal{C}}, \tilde{\nu})$  be ergodic automorphisms of standard probability Borel spaces. Consider the corresponding  $\mathbb{Z}^2$ -action, i.e. the action of  $(\tilde{T}^n \times \tilde{S}^m)_{n,m \in \mathbb{Z}^2}$  on  $(\tilde{X} \times \tilde{Y}, \tilde{\mathcal{B}} \otimes \tilde{\mathcal{C}}, \tilde{\mu} \otimes \tilde{\nu})$ . Suppose that  $\tilde{A}$  is a factor of the  $\mathbb{Z}^2$ -action. Let  $\tilde{B} \otimes \tilde{C}$  be the smallest product factor containing  $\tilde{A}$  <sup>(13)</sup>. Then there exists a compact metric group  $G$  and two continuous 1 – 1 homomorphisms*

$$g \mapsto \tilde{R}_g \in C(\tilde{T}), \quad g \mapsto \tilde{R}'_g \in C(\tilde{S}),$$

such that  $\tilde{A} = \{A \in \tilde{B} \otimes \tilde{C} : (\tilde{R}_g \times \tilde{R}'_g)A = A\}$ .

<sup>(12)</sup> By separability, the smallest  $\sigma$ -algebra including all factors of  $S$  from  $\mathcal{JF}(R)$  is countably generated, whence it corresponds to a joining of at most countably many factors.

<sup>(13)</sup> The existence of such a factor was also proved in [6].



From the above proposition and from (3.6) applied to  $\mathcal{A}_{R \times R}(S \times S)$ , it follows that  $\mathcal{A}_{R \times R}(S \times S)$  is a product- $\sigma$ -algebra. Therefore

$$(3.7) \quad \mathcal{A}_{R \times R}(S \times S) = \mathcal{A}_R(S) \otimes \mathcal{A}_R(S),$$

as this is the largest product  $\sigma$ -algebra in  $\mathcal{JF}(R \times R)$ .

Moreover, using (3.5) and a result from [4] (see also [14]), we conclude that for any  $\lambda \in \mathcal{J}^e(R, S)$  we have  $\lambda = (\lambda|_{\mathcal{D} \otimes \mathcal{A}_R(S)})^\wedge$ , i.e.  $\lambda$  is the relatively independent extension of  $\lambda|_{\mathcal{D} \otimes \mathcal{A}_R(S)}$ . In particular, using (3.7), for  $\lambda \in \mathcal{J}^e(R, S \times S)$  we obtain

$$(3.8) \quad \lambda = (\lambda|_{\mathcal{D} \otimes \mathcal{A}_R(S) \otimes \mathcal{A}_R(S)})^\wedge.$$

We will now show that JP is an intrinsic property.

**PROPOSITION 3.9.** *Let  $R$  be an ergodic automorphism such that in the class  $\mathcal{JF}(R)$  there is some weakly mixing transformation. An automorphism  $R$  has the JP property if and only if for each weakly mixing  $S \in \mathcal{JF}(R)$  and each  $\lambda \in \mathcal{J}^e(R, S \times S)$ , either  $\lambda = (\lambda|_{Z \times Y_1}) \otimes \nu$  or  $\lambda = (\lambda|_{Z \times Y_2}) \otimes \nu$ .*

*Proof.* Fix a weakly mixing automorphism  $S$  and let  $\lambda \in \mathcal{J}^e(T, S \times S)$ . Then (3.8) holds, equivalently,  $\mathcal{D}$  and  $\mathcal{C} \otimes \mathcal{C}$  are relatively independent over  $\mathcal{A}_R(S) \otimes \mathcal{A}_R(S)$ . Take  $f \in L^\infty(Z, \rho)$ ,  $g_i \in L^\infty(Y_i, \nu)$  for  $i = 1, 2$ . Assume that  $\int g_2 = 0$ . We have

$$(3.9) \quad \int f \otimes g_1 \otimes g_2 d\lambda = \int E^\lambda(f \otimes g_1 \otimes g_2 | \mathcal{A}_R(S) \otimes \mathcal{A}_R(S)) d\lambda \\ \stackrel{(3.8)}{=} \int E^\lambda(f) | \mathcal{A}_R(S) \otimes \mathcal{A}_R(S) E^\lambda(g_1 \otimes g_2 | \mathcal{A}_R(S) \otimes \mathcal{A}_R(S)) d\lambda.$$

To distinguish the two  $\sigma$ -algebras  $\mathcal{A}_R(S)$  we will write  $\mathcal{A}_R^1(S)$  for  $\mathcal{A}_R(S) \otimes \{\emptyset, Y\}$  and  $\mathcal{A}_R^2(S)$  for  $\{\emptyset, Y\} \otimes \mathcal{A}_R(S)$ . By the assumption  $\mathcal{D} \otimes \mathcal{A}_R^1(S) \perp \mathcal{A}_R^2(S)$  (up to a permutation of coordinates) with respect to  $\lambda$ , whence the last integral in (3.9) equals

$$\int E^\lambda(f | \mathcal{A}_R^1(S)) E^\nu(g_1 | \mathcal{A}_R^1(S)) E^\nu(g_2 | \mathcal{A}_R^2(S)) d\lambda \\ \stackrel{(*)}{=} \left( \int E^\lambda(f | \mathcal{A}_R^1(S)) E^\nu(g_1 | \mathcal{A}_R^1(S)) d\lambda \right) \cdot \int E^\nu(g_2 | \mathcal{A}_R^2(S)) d\lambda = 0,$$

where  $(*)$  follows by the independence of  $\mathcal{A}_R^1(S)$  and  $\mathcal{A}_R^2(S)$ . ■

**REMARK 3.10.** Notice that if in  $\mathcal{JF}(R)$  there are no weakly mixing automorphisms, then  $R$  is disjoint from all weakly mixing systems (see Theorem 4.10), which implies the JP property.

**REMARK 3.11.** It was shown in [44], using Rokhlin extensions, that there exists a non-weakly mixing automorphism such that the class  $\mathcal{JF}(R)$  includes some weakly mixing systems <sup>(14)</sup>.

<sup>(14)</sup> Recall that the class of JP systems is closed under taking Rokhlin extensions.

**COROLLARY 3.12.** *Let  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  be such that all of its ergodic self-joinings are weakly mixing. Then  $T_1$  has the JP property if and only if  $\mathcal{T}$  does.*

*Proof.* This is immediate by Remarks 3.6, 3.7 and Proposition 3.9. ■

The following problem remains open:

**QUESTION 3.13.** Does the JP property of a weakly mixing flow  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  imply the JP property for  $T_t$ ,  $t \neq 0$ ?

### 3.3. Isometric, distal and relatively weakly mixing extensions

**3.3.1. Isometric extensions.** Suppose that  $T: (X, \mathcal{B}, \mu) \rightarrow (X, \mathcal{B}, \mu)$  is embeddable into some measurable flow  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$ , i.e.  $T = T_1$ . Let  $\{f'_n : n \in \mathbb{N}\} \subset L^2(X, \mu)$  be dense and let  $f_n := f'_n / \|f'_n\|_{L^2(X, \mu)}$  for  $n \in \mathbb{N}$ . Let  $d_{\{f_n\}_{n \in \mathbb{N}}}$  be the metric defined as in (2.4), using  $\{f_n\}_{n \in \mathbb{N}}$ .

**LEMMA 3.14.** *For any  $\varepsilon > 0$  there exists  $\delta > 0$  such that for all  $\xi \in \mathcal{J}_2(T)$  and  $t \in \mathbb{R}$  with  $|t| < \delta$  we have  $\bar{d}_{\{f_n\}_{n \in \mathbb{N}}}(\xi \circ (T_t \times T_t), \xi) < \varepsilon/4$ .*

*Proof.* Fix  $\varepsilon > 0$ . Let  $N \in \mathbb{N}$  be such that  $\sum_{i > N} 1/2^i < \varepsilon/32$ . Let  $\delta > 0$  be such that for  $1 \leq i \leq N$  and  $t \in \mathbb{R}$  such that  $|t| < \delta$  we have  $\|f_i \circ T_t - f_i\|_{L^2(X, \mu)} < \varepsilon/8$ . Then for  $t \in \mathbb{R}$  and  $\xi \in \mathcal{J}_2(T)$  by the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}
d_{\{f_n\}_{n \in \mathbb{N}}}(\xi \circ (T_t \times T_t), \xi) &= \sum_{i, j \geq 1} \frac{1}{2^{i+j}} \left| \int_{X \times X} (f_i \circ T_t \otimes f_j \circ T_t - f_i \otimes f_j) d\xi \right| \\
&\leq \sum_{i, j \geq 1} \frac{1}{2^{i+j}} \left| \int_{X \times X} (f_i \circ T_t \otimes f_j \circ T_t - f_i \circ T_t \otimes f_j) d\xi \right| \\
&\quad + \sum_{i, j \geq 1} \frac{1}{2^{i+j}} \left| \int_{X \times X} (f_i \circ T_t \otimes f_j - f_i \otimes f_j) d\xi \right| \\
&\leq \sum_{i, j \geq 1} \frac{1}{2^{i+j}} \|f_j \circ T_t - f_j\|_{L^2(X, \mu)} + \sum_{i, j \geq 1} \frac{1}{2^{i+j}} \|f_i \circ T_t - f_i\|_{L^2(X, \mu)} \\
&= 2 \left( \sum_{i=1}^N \frac{1}{2^i} \|f_i \circ T_t - f_i\|_{L^2(X, \mu)} + \sum_{i > N} \frac{1}{2^i} \|f_i \circ T_t - f_i\|_{L^2(X, \mu)} \right) \\
&\leq 2 \left( \sum_{i=1}^N \frac{1}{2^i} \cdot \varepsilon/8 + \varepsilon/16 \right) < \varepsilon/4. \quad \blacksquare
\end{aligned}$$

Let  $A_\sigma := \{(m, p, q) \in \mathbb{N}^3 : m \geq 1, q \geq 0, 0 \leq p \leq 2^q - 1, p \text{ is odd}\}$  and  $\sigma: A_\sigma \rightarrow \mathbb{N}$  be a bijection. For  $i \geq 1$  let

$$g_i := f_m \circ T_{p/2^q} \quad \text{where} \quad (m, p, q) = \sigma(i).$$

Then  $\{g_n : n \in \mathbb{N}\} \subset L^2(X, \mathcal{B}, \mu)$  is clearly a dense subset as it includes  $\{f_n : n \in \mathbb{N}\}$ . Let  $\bar{d}$  be the metric on  $\mathcal{J}_2(T)$  defined as in (2.4), but using  $\{g_n\}_{n \in \mathbb{N}}$  instead of  $\{f_n\}_{n \in \mathbb{N}}$ .

LEMMA 3.15. *For any  $\varepsilon > 0$  and any  $q \in \mathbb{N}$  there exists  $C = C(\varepsilon, q)$  such that for all  $0 \leq p \leq 2^q - 1$  and all  $\rho_1, \rho_2 \in \mathcal{J}_2(T)$  we have*

$$\bar{d}(\rho_1 \circ (T_{p/2^q} \times T_{p/2^q}), \rho_2 \circ (T_{p/2^q} \times T_{p/2^q})) \leq C(\varepsilon, q) \cdot \bar{d}(\rho_1, \rho_2) + \varepsilon/2.$$

*Proof.* Fix  $\varepsilon > 0$  and  $q \in \mathbb{N}$ . Let  $0 \leq p \leq 2^q - 1$  and let  $\rho_1, \rho_2 \in \mathcal{J}_2(T)$ . Take  $N \in \mathbb{N}$  such that

$$(3.10) \quad \sum_{i=1}^N \sum_{j>N} \frac{1}{2^{i+j}} + \sum_{i>N} \sum_{j=1}^N \frac{1}{2^{i+j}} < \varepsilon/2.$$

Let  $a_{i,j} := |\int_{X \times X} g_i \otimes g_j d\rho_1 - \int_{X \times X} g_i \otimes g_j d\rho_2|$  for  $i, j \geq 1$ . Then

$$\bar{d}(\rho_1, \rho_2) = \sum_{i,j \geq 1} \frac{1}{2^{i+j}} a_{i,j}.$$

Let  $\pi : \mathbb{N} \rightarrow \mathbb{N}$  be a bijection defined in the following way:

$$\pi(i) = j \Leftrightarrow \begin{cases} \sigma^{-1}(i) = (m, p_i, q_i), \\ \sigma^{-1}(j) = (m, p_j, q_j), \end{cases}$$

for some  $m \in \mathbb{N}$  and  $q_i, q_j \in \mathbb{N}$ ,  $0 \leq p_i \leq 2^{q_i} - 1$ ,  $0 \leq p_j \leq 2^{q_j} - 1$  such that

$$\frac{p_i}{2^{q_i}} + \frac{p}{2^q} - \frac{p_j}{2^{q_j}} \in \mathbb{Z}.$$

The permutation  $\pi$  is well-defined as the addition of  $p/2^q \bmod 1$  is a bijection of the set  $\{p_i/2^{q_i} : (1, p_i, q_i) \in A_\sigma\}$ . Recall also that  $\rho_1, \rho_2$  are  $T \times T$ -invariant as elements of  $\mathcal{J}_2(T)$ . Therefore

$$\bar{d}(\rho_1 \circ (T_{p/2^q} \times T_{p/2^q}), \rho_2 \circ (T_{p/2^q} \times T_{p/2^q})) = \sum_{i,j \geq 1} \frac{1}{2^{i+j}} a_{\pi(i), \pi(j)}.$$

By (3.10) we obtain

$$\begin{aligned} \sum_{i,j \geq 1} \frac{1}{2^{i+j}} a_{\pi(i), \pi(j)} &\leq \sum_{i,j=1}^N \frac{1}{2^{i+j}} a_{\pi(i), \pi(j)} + \varepsilon/2 \\ &= \sum_{i,j=1}^N 2^{\pi(i)-i+\pi(j)-j} \frac{1}{2^{\pi(i)+\pi(j)}} a_{\pi(i), \pi(j)} + \varepsilon/2 \\ &\leq \sum_{i,j=1}^{\max\{\pi(i): 1 \leq i \leq N\}} 2^{2 \max\{|\pi(i)-i|: 1 \leq i \leq N\}} \frac{1}{2^{i+j}} a_{i,j} + \varepsilon/2 \\ &\leq 2^{2 \max\{|\pi(i)-i|: 1 \leq i \leq N\}} \sum_{i,j \geq 1} \frac{1}{2^{i+j}} a_{i,j} + \varepsilon/2 = C(\varepsilon, p, q) \cdot \bar{d}(\rho_1, \rho_2) + \varepsilon/2, \end{aligned}$$

where  $C(\varepsilon, p, q) = 2^{2 \max\{|\pi(i)-i|: 1 \leq i \leq N\}}$ . To end the proof it suffices to put  $C(\varepsilon, q) := \max\{C(\varepsilon, p, q) : 0 \leq p \leq 2^q - 1\}$ . ■

LEMMA 3.16. *For any  $\varepsilon > 0$  there is  $\delta > 0$  such that for any  $\xi_1, \xi_2 \in \mathcal{J}_2(T)$ , if  $\bar{d}(\xi_1, \xi_2) < \delta$  then  $\bar{d}(\xi_1 \circ (T_t \times T_t), \xi_2 \circ (T_t \times T_t)) < \varepsilon$  for all  $t \in \mathbb{R}$ . In particular,  $\bar{d}(\int_0^1 \xi_1 \circ (T_t \times T_t) dt, \int_0^1 \xi_2 \circ (T_t \times T_t) dt) < \varepsilon$ .*

*Proof.* Fix  $\varepsilon > 0$ . By Lemma 3.14, we can find  $\delta_0 > 0$  such that for  $|t| < \delta_0$  and  $\xi \in \mathcal{J}_2(T)$ ,

$$(3.11) \quad \bar{d}(\xi \circ (T_t \times T_t), \xi) < \varepsilon/8.$$

Let  $q \in \mathbb{N}$  be such that  $1/2^q < \delta_0$  and take  $C = C(\varepsilon, q_0)$  as in Lemma 3.15. For any  $t \in \mathbb{R}$  find  $0 \leq p \leq 2^q - 1$  such that  $|t - p/2^q| < \delta_0$ . Using (3.11) we obtain

$$\begin{aligned} \bar{d}(\xi_1 \circ (T_t \times T_t), \xi_2 \circ (T_t \times T_t)) \\ \leq \bar{d}(\xi_1 \circ (T_{p/2^q} \times T_{p/2^q}), \xi_2 \circ (T_{p/2^q} \times T_{p/2^q})) + \varepsilon/4 \\ \leq C\bar{d}(\xi_1, \xi_2) + \varepsilon/4 + \varepsilon/4 = C\bar{d}(\xi_1, \xi_2) + \varepsilon/2, \end{aligned}$$

Let  $0 < \delta < \varepsilon/(2C)$ . Then  $\bar{d}(\xi_1, \xi_2) < \delta$  implies  $\bar{d}(\xi_1 \circ (T_t \times T_t), \xi_2 \circ (T_t \times T_t)) < \varepsilon$ . It follows that

$$\begin{aligned} \bar{d}\left(\int_0^1 \xi_1 \circ (T_t \times T_t) dt, \int_0^1 \xi_2 \circ (T_t \times T_t) dt\right) \\ \leq \int_0^1 \bar{d}(\xi_1 \circ (T_t \times T_t), \xi_2 \circ (T_t \times T_t)) dt < \varepsilon. \quad \blacksquare \end{aligned}$$

PROPOSITION 3.17. *Let  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  be an ergodic flow on  $(X, \mathcal{B}, \underline{\nu})$  with a factor  $\mathcal{A}$ . Let  $\underline{\nu} = \underline{\tilde{\nu}}|_{\mathcal{A}}$ . Let  $\underline{\tilde{\nu}} = \int_0^1 \tilde{\nu} \circ T_t dt$  and  $\underline{\nu} = \int_0^1 \nu \circ T_t|_{\mathcal{A}} dt$  for some  $\tilde{\nu} \in \mathcal{P}_{\mathbb{Z}}^e(T_1), \nu \in \mathcal{P}_{\mathbb{Z}}^e(T_1|_{\mathcal{A}})$  <sup>(15)</sup>. Additionally assume that  $\tilde{\nu}|_{\mathcal{A}} = \nu$ . Then the extension  $(T_1, \tilde{\nu}) \rightarrow (T_1|_{\mathcal{A}}, \nu)$  is isometric if and only if  $(\mathcal{T}, \underline{\tilde{\nu}}) \rightarrow (\mathcal{T}|_{\mathcal{A}}, \underline{\nu})$  is isometric.*

*Proof.* Suppose that  $(\mathcal{T}, \underline{\tilde{\nu}}) \rightarrow (\mathcal{T}|_{\mathcal{A}}, \underline{\nu})$  is isometric. Let  $\overline{\mathcal{T}} = (\overline{T}_t)_{t \in \mathbb{R}} : (\overline{X}, \overline{\mathcal{B}}, \overline{\underline{\nu}}) \rightarrow (\overline{X}, \overline{\mathcal{B}}, \overline{\underline{\nu}})$  be an ergodic extension of  $\mathcal{T}$  such that  $(\overline{\mathcal{T}}, \overline{\underline{\nu}}) \rightarrow (\mathcal{T}|_{\mathcal{A}}, \underline{\nu})$  is a compact group extension. By Lemma 2.2 there exists  $\overline{\nu} \in \mathcal{P}_{\mathbb{Z}}^e(\overline{T}_1)$  such that  $\overline{\underline{\nu}} = \int_0^1 \overline{\nu} \circ \overline{T}_t dt$  (it may happen that  $\overline{\underline{\nu}} = \overline{\nu}$ ). By the uniqueness of ergodic decomposition and by Lemma 2.3, there exist  $k, l, m \in$

<sup>(15)</sup> It may happen that  $\underline{\nu} = \nu$  or  $\underline{\tilde{\nu}} = \tilde{\nu}$ . The existence of  $\nu$  and  $\tilde{\nu}$  follows from Lemma 2.2.

$\mathbb{N} \cup \{\infty\}$  such that <sup>(16)</sup>

$$\underline{\nu} = k \int_0^{1/k} \bar{\nu} \circ \bar{T}_t dt, \quad \tilde{\nu} = kl \int_0^{1/kl} \tilde{\nu} \circ T_t dt, \quad \underline{\nu} = klm \int_0^{1/klm} \nu \circ T_t|_{\mathcal{A}} dt.$$

We may assume that  $\bar{\nu}|_{\mathcal{B}} = \tilde{\nu}$ . Since  $(\bar{\mathcal{T}}, \underline{\nu}) \rightarrow (\mathcal{T}|_{\mathcal{A}}, \underline{\nu})$  is a compact group extension, so is  $(\bar{T}_1, \underline{\nu}) \rightarrow (T_1|_{\mathcal{A}}, \underline{\nu})$ . Notice that in the decomposition  $\underline{\nu} = k \int_0^{1/k} \bar{\nu} \circ \bar{T}_t dt$  the only measures which project down to  $\nu$  are of the form  $\bar{\nu} \circ T_{i/(lm)}$  for  $0 \leq i \leq lm - 1$ . Therefore

$$\left( \bar{T}_1, \frac{1}{lm} \sum_{i=0}^{lm-1} \bar{\nu} \circ T_{i/(lm)} \right) \rightarrow (T_1|_{\mathcal{A}}, \nu)$$

is a compact group extension (if  $lm = \infty$ , we interpret  $\frac{1}{lm} \sum_{i=0}^{lm-1} \bar{\nu} \circ T_{i/(lm)}$  as  $\underline{\nu}$ ). It follows that also  $(\bar{T}_1, \bar{\nu}) \rightarrow (T_1, \nu)$  is a compact group extension. Thus  $(T_1, \tilde{\nu}) \rightarrow (T_1, \nu)$  is isometric, which ends the first part of the proof.

Assume now that  $(T_1, \tilde{\nu}) \rightarrow (T_1|_{\mathcal{A}}, \nu)$  is isometric. As above, there exist  $k, l \in \mathbb{N} \cup \{\infty\}$  such that

$$(3.12) \quad \tilde{\nu} = k \int_0^{1/k} \tilde{\nu} \circ T_t dt, \quad \underline{\nu} = kl \int_0^{1/kl} \nu \circ T_t|_{\mathcal{A}} dt$$

are the corresponding ergodic decompositions (see footnotes 15 and 16). By Proposition 2.11 the extension  $(T_1, \tilde{\nu}) \rightarrow (T_1|_{\mathcal{A}}, \nu)$  is T-compact. Fix  $\varepsilon > 0$ . Let  $0 < \delta < \varepsilon$  be as in Lemma 3.14. Let  $N \in \mathbb{N}$  come from the definition of T-compactness for  $(T_1, \tilde{\nu}) \rightarrow (T_1|_{\mathcal{A}}, \nu)$  with  $\varepsilon$  replaced with  $\delta$ . Take  $\underline{\lambda} \in J_N^e((T, \tilde{\nu}); \mathcal{A})$ . Then  $\underline{\lambda} \in \mathcal{P}_{\mathbb{Z}}(T_1^{\times N})$  and either (i)  $\underline{\lambda} \in \mathcal{P}_{\mathbb{Z}}^e(T_1^{\times N})$ , or (ii)  $\underline{\lambda} \notin \mathcal{P}_{\mathbb{Z}}^e(T_1^{\times N})$ .

Case (i) may occur only if  $k = \infty$ , i.e.  $\tilde{\nu} = \nu$ ,  $\underline{\nu} = \nu$  (otherwise the marginals of  $\underline{\lambda}$  are not ergodic for  $T_1$ , so  $\underline{\lambda}$  is not ergodic for  $T_1^{\times N}$ ). Then by Proposition 2.11 there exist  $1 \leq i < j \leq N$  such that

$$\bar{d}(\underline{\lambda}|_{X_i \times X_j}, \Delta_{X_i \times X_j}(\tilde{\nu})) < \delta < \varepsilon.$$

For case (ii), to fix ideas, we will assume that  $k, l \in \mathbb{N}$ , i.e.  $\tilde{\nu} \neq \nu$  and  $\underline{\nu} \neq \nu$  (in the other cases the proof is similar). By Lemma 2.2 there exists  $\lambda \in \mathcal{P}_{\mathbb{Z}}^e(T_1^{\times N})$  such that

$$(3.13) \quad \underline{\lambda} = \int_0^1 \lambda \circ T_t^{\times N} dt.$$

---

<sup>(16)</sup> We interpret  $\int_0^{1/\infty} \lambda \circ S_t dt$  as  $\underline{\lambda}$  for any measure  $\lambda$  and any flow  $(S_t)_{t \in \mathbb{R}}$  such that this formula makes sense.

We will show that

$$(3.14) \quad \lambda \text{ can be chosen from } \mathcal{J}_N^e((T_1, \tilde{\nu}); \mathcal{A}).$$

We will use the fact that ergodic decompositions are unique. We have

$$(3.15) \quad \begin{aligned} \underline{\lambda}|_{\mathcal{A}^{\otimes N}} &= \int_0^1 \lambda|_{\mathcal{A}^{\otimes N}} \circ (T_t|_{\mathcal{A}})^{\times N} dt, \\ \underline{\lambda}|_{\mathcal{A}^{\otimes N}} &= \Delta_{\mathcal{A}^{\otimes N}}(\underline{\nu}) = kl \int_0^{1/(kl)} \Delta_{\mathcal{A}^{\otimes N}}(\nu) \circ (T_t|_{\mathcal{A}})^{\times N} dt \end{aligned}$$

(the latter follows from the second formula in (3.12)). Therefore we may assume (changing  $\lambda \in \mathcal{P}_{\mathbb{Z}}^e(T_1^{\times N})$  if necessary) that

$$(3.16) \quad \lambda|_{\mathcal{A}^{\otimes N}} = \Delta_{\mathcal{A}^{\otimes N}}(\nu)$$

and  $\underline{\lambda}|_{\mathcal{A}^{\otimes N}} = kl \int_0^{1/(kl)} \lambda|_{\mathcal{A}^{\otimes N}} \circ (T_t|_{\mathcal{A}})^{\times N} dt$ . Moreover, for  $1 \leq i \leq N$  we have

$$\underline{\lambda}|_{X_i} = k \int_0^{1/k} \lambda|_{X_i} \circ T_t dt, \quad \underline{\lambda}|_{X_i} = \tilde{\nu} = k \int_0^{1/k} \tilde{\nu} \circ T_t dt,$$

whence there exists  $t_i \in [0, 1/k)$  such that

$$(3.17) \quad \lambda|_{X_i} = \tilde{\nu} \circ T_{t_i}.$$

It follows that

$$(\lambda|_{X_i})|_{\mathcal{A}} = \tilde{\nu}|_{\mathcal{A}} \circ T_{t_i}|_{\mathcal{A}} = \nu \circ (T_{t_i}|_{\mathcal{A}}).$$

On the other hand, by (3.16), we have  $(\lambda|_{X_i})|_{\mathcal{A}} = \nu$ . Hence  $t_i = 0$ , and (3.14) follows from (3.16) and (3.17). By Proposition 2.11 there exist  $1 \leq i < j \leq N$  such that

$$(3.18) \quad \bar{d}(\lambda|_{X_i \times X_j}, \Delta_{X_i \times X_j}(\tilde{\nu})) < \delta.$$

By (3.13) we have

$$(3.19) \quad \underline{\lambda}|_{X_i \times X_j} = \int_0^1 \lambda|_{X_i \times X_j} \circ (T_t \times T_t) dt.$$

Moreover,

$$(3.20) \quad \Delta_{X_i \times X_j}(\underline{\nu}) = \int_0^1 \Delta_{X_i \times X_j}(\tilde{\nu}) \circ (T_t \times T_t) dt$$

From the choice of  $\delta$ , from (3.18)–(3.20) and from Lemma 3.16, we deduce that  $\bar{d}(\underline{\lambda}|_{X_i \times X_j}, \Delta_{X_i \times X_j}(\underline{\nu})) < \varepsilon$ . The claim follows from Proposition 2.11. ■

**PROPOSITION 3.18.** *Let  $l \in \mathbb{N}$  and let  $T$  be an ergodic automorphism of  $(X, \mathcal{B}, \tilde{\nu})$  with a factor  $\mathcal{A}$ . Let  $\underline{\nu} = \tilde{\nu}|_{\mathcal{A}}$ . Let  $\underline{\nu} = (1/l) \sum_{i=1}^{l-1} \tilde{\nu} \circ T^i$  and  $\underline{\nu} = (1/l) \sum_{i=0}^{l-1} \nu \circ T^i|_{\mathcal{A}}$  for some  $\tilde{\nu} \in \mathcal{P}_{\mathbb{Z}}^e(T^l), \nu \in \mathcal{P}_{\mathbb{Z}}^e(T^l|_{\mathcal{A}})$ . Additionally*

assume that  $\tilde{\nu}|_{\mathcal{A}} = \nu$ . Then the extension  $(T^l, \tilde{\nu}) \rightarrow (T^l|_{\mathcal{A}}, \nu)$  is isometric if and only if  $(T, \tilde{\underline{\nu}}) \rightarrow (T|_{\mathcal{A}}, \underline{\nu})$  is isometric.

*Proof.* The proof goes along the same lines as the proof of Proposition 3.17. ■

### 3.3.2. Distal extensions and Furstenberg–Zimmer decomposition

*Distal extensions via isometric extensions.* We start with

**PROPOSITION 3.19.** *Let  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  be an ergodic flow on  $(X, \mathcal{B}, \tilde{\underline{\nu}})$  with a factor  $\mathcal{A}$ . Let  $\underline{\nu} = \tilde{\underline{\nu}}|_{\mathcal{A}}$ . Let  $\tilde{\underline{\nu}} = \int_0^1 \tilde{\nu} \circ T_t dt$  and  $\underline{\nu} = \int_0^1 \nu \circ T_t dt$  for some  $\tilde{\nu} \in \mathcal{P}_{\mathbb{Z}}^e(T_1)$ ,  $\nu \in \mathcal{P}_{\mathbb{Z}}^e(T_1|_{\mathcal{A}})$  <sup>(17)</sup>. Additionally assume that  $\tilde{\nu}|_{\mathcal{A}} = \nu$ . Then the extension  $(T_1, \tilde{\nu}) \rightarrow (T_1|_{\mathcal{A}}, \nu)$  is distal if and only if  $(\mathcal{T}, \tilde{\underline{\nu}}) \rightarrow (\mathcal{T}|_{\mathcal{A}}, \underline{\nu})$  is distal.*

*Proof.* Suppose that  $(\mathcal{T}, \tilde{\underline{\nu}}) \rightarrow (\mathcal{T}|_{\mathcal{A}}, \underline{\nu})$  is distal. Since any factor of  $\mathcal{T}$  is a factor of  $T_1$ , it clearly follows from the definition of a distal extension and Proposition 3.17 that also  $(T_1, \tilde{\nu}) \rightarrow (T_1|_{\mathcal{A}}, \nu)$  is distal.

Suppose now that  $(T_1, \tilde{\nu}) \rightarrow (T_1|_{\mathcal{A}}, \nu)$  is distal. Let  $(\mathcal{C}_\alpha)_{\alpha \leq \beta}$  be a transfinite sequence of factors of  $T_1$  such that  $\mathcal{C}_0 = \mathcal{A}$ ,  $\mathcal{C}_\beta = \mathcal{B}$ , the extension  $\mathcal{C}_{\alpha+1} \rightarrow \mathcal{C}_\alpha$  for  $T_1$  is isometric and for  $\alpha$  being a limit ordinal,  $\mathcal{C}_\alpha$  is the inverse limit of the preceding factors. We may assume that for each  $\alpha$  the factor  $\mathcal{C}_{\alpha+1}$  is the largest extension of  $\mathcal{C}_\alpha$  inside  $\mathcal{B}$  such that  $\mathcal{C}_{\alpha+1} \rightarrow \mathcal{C}_\alpha$  is isometric. In particular, for each  $\alpha$ ,  $\mathcal{C}_\alpha$  is  $\mathcal{T}$ -invariant, i.e. it is a factor of  $\mathcal{T}$ . Indeed, for  $\alpha$  which is not a limit ordinal,  $T_1|_{\mathcal{C}_\alpha}$  is isomorphic to  $T_1|_{T_t \mathcal{C}_\alpha}$ , and both  $\mathcal{C}_\alpha \rightarrow \mathcal{C}_{\alpha-1}$ ,  $T_t \mathcal{C}_\alpha \rightarrow \mathcal{C}_{\alpha-1}$  are isometric, whence  $\mathcal{C}_\alpha \vee T_t \mathcal{C}_\alpha \rightarrow \mathcal{C}_{\alpha-1}$  is also isometric, which follows e.g. from (2.7) <sup>(18)</sup>. For  $\alpha$  which is a limit ordinal it suffices to notice that the property of being  $\mathcal{T}$ -invariant is preserved under taking inverse limits. ■

The same proof works in the co-finite case:

**PROPOSITION 3.20.** *Let  $T$  be an ergodic automorphism acting on  $(X, \mathcal{B}, \tilde{\underline{\nu}})$  with a factor  $\mathcal{A}$ . Fix  $l \in \mathbb{N}$ . Let  $\underline{\nu} = \tilde{\underline{\nu}}|_{\mathcal{A}}$ . Let  $\tilde{\underline{\nu}} = (1/l) \sum_{i=0}^{l-1} \tilde{\nu} \circ T_i$  and  $\underline{\nu} = (1/l) \sum_{i=0}^{l-1} \nu \circ T_i|_{\mathcal{A}}$  for some  $\tilde{\nu} \in \mathcal{P}_{\mathbb{Z}}^e(T^l)$ ,  $\nu \in \mathcal{P}_{\mathbb{Z}}^e(T^l|_{\mathcal{A}})$ . Additionally assume that  $\tilde{\nu}|_{\mathcal{A}} = \nu$ . Then the extension  $(T^l, \tilde{\nu}) \rightarrow (T^l|_{\mathcal{A}}, \nu)$  is distal if and only if  $(T, \tilde{\underline{\nu}}) \rightarrow (T|_{\mathcal{A}}, \underline{\nu})$  is distal.*

*Furstenberg–Zimmer decomposition.* Using the techniques from the proof of Proposition 3.19 and arguing by contradiction one can show more:

**COROLLARY 3.21.** *Under the assumptions of Proposition 3.19 the following statements are true and equivalent:*

<sup>(17)</sup> See footnote 16.

<sup>(18)</sup>  $\mathcal{C}_\alpha \vee T_t \mathcal{C}_\alpha$  stands for the  $\sigma$ -algebra generated by  $\mathcal{C}_\alpha$  and  $T_t \mathcal{C}_\alpha$ .

- (i)  $(T_1, \tilde{\nu}) \rightarrow (T_1|_{\mathcal{A}}, \nu)$  is distal if and only if  $(\mathcal{T}, \tilde{\underline{\nu}}) \rightarrow (\mathcal{T}|_{\mathcal{A}}, \underline{\nu})$  is distal.
- (ii)  $(T_1, \tilde{\nu}) \rightarrow (T_1|_{\mathcal{A}}, \nu)$  is relatively weakly mixing if and only if  $(\mathcal{T}, \tilde{\underline{\nu}}) \rightarrow (\mathcal{T}|_{\mathcal{A}}, \underline{\nu})$  is relatively weakly mixing.
- (iii) The Furstenberg–Zimmer decompositions for  $(T_1, \tilde{\nu}) \rightarrow (T_1|_{\mathcal{A}}, \nu)$  and for  $(\mathcal{T}, \tilde{\underline{\nu}}) \rightarrow (\mathcal{T}|_{\mathcal{A}}, \underline{\nu})$  are the same, i.e. there exists  $\mathcal{A} \subset \mathcal{C} \subset \mathcal{B}$  which is a factor of  $\mathcal{T}$  such that  $\mathcal{B} \rightarrow \mathcal{C}$  is relatively weakly mixing for both  $(\mathcal{T}, \tilde{\underline{\nu}})$  and  $(T_1, \tilde{\nu})$ , and  $\mathcal{C} \rightarrow \mathcal{A}$  is distal for both  $(\mathcal{T}, \tilde{\underline{\nu}}|_{\mathcal{C}})$  and  $(T_1, \tilde{\nu}|_{\mathcal{C}})$ .

In the co-finite case we have the following:

**COROLLARY 3.22.** *Under the assumptions of Proposition 3.20 the following statements are true and equivalent:*

- (i)  $(T^l, \tilde{\nu}) \rightarrow (T^l|_{\mathcal{A}}, \nu)$  is distal if and only if  $(T, \tilde{\underline{\nu}}) \rightarrow (T|_{\mathcal{A}}, \underline{\nu})$  is distal.
- (ii)  $(T^l, \tilde{\nu}) \rightarrow (T^l|_{\mathcal{A}}, \nu)$  is relatively weakly mixing if and only if  $(T, \tilde{\underline{\nu}}) \rightarrow (T|_{\mathcal{A}}, \underline{\nu})$  is relatively weakly mixing.
- (iii) The Furstenberg–Zimmer decompositions for  $(T^l, \tilde{\nu}) \rightarrow (T^l|_{\mathcal{A}}, \nu)$  and for  $(T, \tilde{\underline{\nu}}) \rightarrow (T|_{\mathcal{A}}, \underline{\nu})$  are the same, i.e. there exists  $\mathcal{A} \subset \mathcal{C} \subset \mathcal{B}$  which is a factor of  $T$  such that  $\mathcal{B} \rightarrow \mathcal{C}$  is relatively weakly mixing for both  $(T, \tilde{\underline{\nu}})$  and  $(T^l, \tilde{\nu})$  and  $\mathcal{C} \rightarrow \mathcal{A}$  is distal for both  $(T, \tilde{\underline{\nu}}|_{\mathcal{C}})$  and  $(T^l, \tilde{\nu}|_{\mathcal{C}})$ .

**3.4. Quasi- and distal simplicity.** In this section we prove counterparts of Propositions 3.1 and 3.5 for the 2-QS and 2-DS properties.

### 3.4.1. Quasi-simplicity

**PROPOSITION 3.23.** *Let  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  be a measure-preserving flow on  $(X, \mathcal{B}, \mu)$  with  $T_1$  ergodic. Then  $\mathcal{T}$  is 2-QS if and only if  $T_1$  is 2-QS.*

*Proof.* Suppose that  $\mathcal{T}$  is 2-QS and let  $\lambda \in \mathcal{J}_2^e(T_1)$ ,  $\lambda \neq \mu \otimes \mu$ . There are two possibilities:

- (i)  $\lambda = \lambda \circ (T_t \times T_t)$  for all  $t \in \mathbb{R}$ ,
- (ii)  $\lambda \perp \lambda \circ (T_t \times T_t)$  for some  $t \in (0, 1)$ .

In case (i) apply Proposition 3.17 to  $(\mathcal{T} \times \mathcal{T}, \lambda)$ ,  $(T_1 \times T_1, \lambda)$  and the coordinate factors. In case (ii) consider  $\underline{\lambda} := \int_0^1 \lambda \circ (T_t \times T_t) dt$  and apply Proposition 3.17 to  $(\mathcal{T} \times \mathcal{T}, \underline{\lambda})$ ,  $(T_1 \times T_1, \lambda)$  and the coordinate factors.

Suppose now that  $T_1$  is 2-QS. Let  $\underline{\lambda} \in \mathcal{J}_2^e(\mathcal{T})$ ,  $\lambda \neq \mu \otimes \mu$ . Again there are two possibilities:

- (i)  $\underline{\lambda} \in \mathcal{J}_2^e(\mathcal{T})$ ,
- (ii)  $\underline{\lambda} \notin \mathcal{J}^e(\mathcal{T})$ .



In case (i) apply Proposition 3.17 to  $(\mathcal{T} \times \mathcal{T}, \underline{\lambda})$ ,  $(T_1 \times T_1, \lambda)$  and the coordinate factors. In case (ii) take  $\lambda \in \mathcal{J}_2^e(T_1)$  such that  $\underline{\lambda} = \int_0^1 \lambda \circ (T_t \times T_t) dt$  and apply Proposition 3.17 to  $(\mathcal{T} \times \mathcal{T}, \underline{\lambda})$ ,  $(T_1 \times T_1, \lambda)$  and the coordinate factors. ■

**PROPOSITION 3.24.** *Let  $l \geq 1$  and let  $T$  be an automorphism of  $(X, \mathcal{B}, \mu)$  such that  $T^l$  is ergodic. Then  $T$  is 2-QS if and only if  $T^l$  is 2-QS.*

*Proof.* The proof goes along the same lines as the proof of Proposition 3.23. It suffices to replace integrals by finite averages and apply Proposition 3.18 instead of Proposition 3.17. ■

**REMARK 3.25.** In the extended version of the paper (available on arXiv) we also give a short proof that, given  $\mathcal{T}$  such that  $T_1$  is ergodic,  $\mathcal{T}$  is 2-QS whenever  $T_1$  is 2-fold simple.

### 3.4.2. Distal simplicity

**PROPOSITION 3.26.** *Let  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  be a measure-preserving flow on  $(X, \mathcal{B}, \mu)$  with  $T_1$  ergodic. Then  $\mathcal{T}$  is 2-DS if and only if  $T_1$  is 2-DS.*

*Proof.* In the proof of Proposition 3.23 replace “Proposition 3.17” with “Proposition 3.19”. ■

**PROPOSITION 3.27.** *Let  $l \geq 1$  and let  $T$  be an automorphism of  $(X, \mathcal{B}, \mu)$  such that  $T^l$  is ergodic. Then  $T$  is 2-DS if and only if  $T^l$  is 2-DS.*

*Proof.* In the proof of Proposition 3.24 replace “Proposition 3.18” with “Proposition 3.20”. ■

## 4. (Non-)uniqueness of embedding

**4.1. Uniqueness.** We will now discuss the problem of the uniqueness of embedding of automorphisms into measurable flows (and the uniqueness of roots of automorphisms) from the point of view of self-joining properties.

**REMARK 4.1.** Given a flow  $\mathcal{T}$  the following conditions are equivalent:

- (i) for any flow  $\mathcal{S}$  the condition  $T_1 = S_1$  implies  $\mathcal{T} = \mathcal{S}$ ,
- (ii) for any flow  $\mathcal{S}$  such that  $T_1 \simeq S_1$ , the isomorphism between  $T_1$  and  $S_1$  is an isomorphism between  $\mathcal{T}$  and  $\mathcal{S}$ .

**REMARK 4.2.** For weakly mixing flows  $\mathcal{T}, \mathcal{S}$  such that  $T_1 \not\sim S_1$ , we always have  $\mathcal{T} \not\sim \mathcal{S}$ . Indeed, every non-trivial joining between  $T_1$  and  $S_1$  lifts to a non-trivial joining between  $\mathcal{T}$  and  $\mathcal{S}$ . This holds in particular when  $T_1 = S_1$ .

**PROPOSITION 4.3.** *Let  $\mathcal{T}$  be a weakly mixing, measure-preserving flow on  $(X, \mathcal{B}, \mu)$  with the MSJ property. Then  $T_1$  has a unique embedding.*

*Proof.* Let  $\mathcal{S} = (S_t)_{t \in \mathbb{R}}$  on  $(X, \mathcal{B}, \mu)$  be such that  $T_1 = S_1$ . Then  $T_1 = S_1$  is 2-fold simple by Proposition 3.1. It follows from Proposition 3.23 that  $\mathcal{S}$  is

2-QS. Since, by Remark 4.2,  $\mathcal{T} \not\sim \mathcal{S}$ , we find that  $\mathcal{T}$  and  $\mathcal{S}$  have a non-trivial common factor (see [19]). Since  $\mathcal{T}$  has the MSJ property this factor is equal to  $\mathcal{B}$  and there exists  $\mathcal{C} \subset \mathcal{B}$  invariant under  $\mathcal{S}$  and such that  $\mathcal{S}|_{\mathcal{C}} = \mathcal{T}$ . On the other hand,  $T_1 = S_1$ , which implies  $\mathcal{C} = \mathcal{B}$  and the claim follows. ■

**PROPOSITION 4.4.** *If  $T$  is a weakly mixing automorphism with the MSJ property then for  $k \in \mathbb{N}$  the automorphism  $T^k$  has a unique root of order  $k$ .*

*Proof.* The proof of Proposition 4.3 can be easily adjusted to the case of automorphisms. ■

We will now show a more universal way of proving uniqueness of embedding. It will be the main tool for proving the analogue of Proposition 4.3 for flows with the R-property.

**PROPOSITION 4.5.** *Let  $\mathcal{T}$  be a weakly mixing flow on  $(X, \mathcal{B}, \mu)$  such that for any weakly mixing flow  $\mathcal{S}$  on  $(Y, \mathcal{C}, \nu)$  and any  $\lambda \in \mathcal{J}^e(\mathcal{T}, \mathcal{S})$  either  $\lambda = \mu \otimes \nu$  or the fibers in the extension  $(\mathcal{T} \times \mathcal{S}, \lambda) \rightarrow (\mathcal{S}, \nu)$  are finite. Then for any  $\lambda \in \mathcal{J}^e(\mathcal{T}, \mathcal{S})$ , the flow  $(\mathcal{T} \times \mathcal{S}, \lambda)$  is weakly mixing.*

*Proof.* Let  $\lambda \in \mathcal{J}^e(\mathcal{T}, \mathcal{S})$ . If  $\lambda = \mu \otimes \nu$  then  $(\mathcal{T} \times \mathcal{S}, \lambda)$  is clearly weakly mixing. Suppose that  $\lambda \neq \mu \otimes \nu$  and  $(\mathcal{T} \times \mathcal{S}, \lambda)$  is not weakly mixing. Denote by  $\mathcal{K}$  the Kronecker factor of  $(\mathcal{T} \times \mathcal{S}, \lambda)$ . Then  $(\mathcal{T} \times \mathcal{S}, \lambda)$  has  $\mathcal{B} \otimes \mathcal{K}$  as a factor <sup>(19)</sup>. Consider the following disintegrations:

$$\lambda = \int \mu_x d\mu, \quad \lambda = \int \lambda_\omega d\lambda|_{\mathcal{B} \otimes \mathcal{K}}, \quad \lambda|_{\mathcal{B} \otimes \mathcal{K}} = \int \tilde{\lambda}_z d\mu.$$

For  $f \in L^1(X \times Y, \mathcal{B} \otimes \mathcal{C}, \lambda)$  we have  $\int f d\lambda = \int \int f d\mu_x d\mu$ . On the other hand,  $\int f d\lambda = \int \int f d\lambda_\omega d\lambda|_{\mathcal{B} \otimes \mathcal{K}} = \int \int \int f d\lambda_\omega d\tilde{\lambda}_z d\mu = \int \int \int f d\tilde{\lambda}_z d\lambda_\omega d\mu$ . Since  $(\mathcal{T} \times \mathcal{S}, \lambda)$  is not weakly mixing, the discrete part of the spectrum forms a countable (infinite) subgroup of  $\mathbb{R}$  and the corresponding factor acts by rotations on the dual group to this group. The measures  $\tilde{\lambda}_z$  are equal to the Haar measure on this group, in particular they have infinite supports. It follows from the uniqueness of disintegration that also the supports of the measures  $\mu_x$  have an infinite number of points. This is however impossible by our assumption that  $\lambda$  is a finite extension of  $\mu$ , i.e. the measures  $\mu_x$  are discrete with a finite number of atoms. ■

**REMARK 4.6.** The above proposition can be strengthened, using the same arguments: If  $\mathcal{T}$  is a weakly mixing flow such that for any weakly mixing flow  $\mathcal{S}$  and any  $\lambda \in \mathcal{J}^e(\mathcal{T}, \mathcal{S})$  the flow  $(\mathcal{T} \times \mathcal{S}, \lambda)$  is also weakly mixing, then any finite group extension of  $\mathcal{T}$  has the same property.

---

<sup>(19)</sup> The symbol  $\otimes$  used here does not mean that  $\mathcal{K} \subset \mathcal{C}$ . By  $\mathcal{B} \otimes \mathcal{K}$  we denote the sub- $\sigma$ -algebra of  $\mathcal{B} \otimes \mathcal{C}$  generated by  $\mathcal{B}$  and by the Kronecker factor  $\mathcal{K}$ . The action of  $\mathcal{T} \times \mathcal{S}$  restricted to  $\mathcal{B} \otimes \mathcal{K}$  is isomorphic to the Cartesian product of  $\mathcal{T}$  and  $(\mathcal{T} \times \mathcal{S})|_{\mathcal{K}}$ .

PROPOSITION 4.7. *Let  $\mathcal{T}$  be a weakly mixing flow with the following property: for any weakly mixing flow  $\mathcal{S}$  and any  $\lambda \in \mathcal{J}^e(\mathcal{T}, \mathcal{S})$  the flow  $(\mathcal{T} \times \mathcal{S}, \lambda)$  is weakly mixing. Then  $T_1$  has a unique embedding.*

*Proof.* Suppose that there exists  $\mathcal{S}$  such that  $T_1 = S_1$  and consider the diagonal joining of  $\Delta \in \mathcal{J}^e(T_1, S_1)$ . If  $\Delta \notin \mathcal{J}(\mathcal{T}, \mathcal{S})$  then, for some  $k \in \mathbb{N}$ ,  $\underline{\Delta} := \int_0^{1/k} \Delta \circ (T_t \times S_t) dt \in \mathcal{J}^e(\mathcal{T}, \mathcal{S})$ , where the measures  $\Delta \circ (T_t \times S_t)$  are pairwise orthogonal for  $t \in [0, 1/k)$ . It follows from Remark 2.5 that  $(\mathcal{T} \times \mathcal{S}, \underline{\Delta})$  is not weakly mixing, which contradicts our assumptions. Hence  $\Delta \in \mathcal{J}(\mathcal{T}, \mathcal{S})$ . ■

By combining Proposition 4.5 with Propositions 4.7 and 2.10, we obtain the following:

COROLLARY 4.8. *Let  $\mathcal{T}$  be a weakly mixing flow with the R-property. Then  $T_1$  has a unique embedding.*

PROPOSITION 4.9. *Let  $\mathcal{T}$  be a weakly mixing flow with the MSJ property. Then for any weakly mixing flow  $\mathcal{S}$  and any  $\lambda \in \mathcal{J}^e(\mathcal{T}, \mathcal{S})$  other than the product measure the extension  $(\mathcal{T} \times \mathcal{S}, \lambda) \rightarrow (\mathcal{S}, \nu)$  has finite fibers.*

Before we begin the proof, let us recall a result on the absence of disjointness and state a necessary lemma.

THEOREM 4.10 ([46]). *Let  $\mathcal{T}$  and  $\mathcal{S}$  be ergodic flows. If  $\mathcal{T}$  and  $\mathcal{S}$  are not disjoint then there exists  $\lambda \in \mathcal{J}_\infty^e(\mathcal{T})$  such that  $(\mathcal{T}^{\times\infty}, \lambda)$  and  $\mathcal{S}$  have a common factor.*

LEMMA 4.11. *Let  $\mathcal{T}, \tilde{\mathcal{T}}, \mathcal{R}$  be ergodic flows. Suppose that the extension  $\tilde{\mathcal{T}} \rightarrow \mathcal{T}$  is isometric (“ $k : 1$ ”). Let  $\tilde{\rho} \in \mathcal{J}^e(\tilde{\mathcal{T}}, \mathcal{R})$  and let  $\rho$  be the restriction of  $\tilde{\rho}$  to a joining between  $\mathcal{T}$  and  $\mathcal{R}$ . Then also the extension  $(\tilde{\mathcal{T}} \times \mathcal{R}, \tilde{\rho}) \rightarrow (\mathcal{T} \times \mathcal{R}, \rho)$  is isometric (“ $k : 1$ ”).*

*Proof.* It suffices to show that the claim remains true if we replace “isometric extension” with “compact group extension”. This is however clear, as for any compact group extension  $\mathcal{T}_\varphi$ , we can write  $\mathcal{T}_\varphi \vee \mathcal{R}$  as  $(\mathcal{T} \vee \mathcal{R})_{\tilde{\varphi}}$ , where  $\tilde{\varphi}(x, y) = \varphi(x)$ . In this way we obtain the first part of the claim (about isometric extensions). The more detailed part (about the extensions being “ $k : 1$ ”) follows from the fact that the group  $G$  where  $\varphi$  and  $\tilde{\varphi}$  take their values is the same. Moreover, any isometric extension with finite fibers is an intermediate extension for some finite group extension. ■

*Proof of Proposition 4.9.* Take  $\mu \otimes \nu \neq \lambda \in \mathcal{J}^e(\mathcal{T}, \mathcal{S})$ . It follows from [20] that there exists  $n \in \mathbb{N}$  such that  $\mathcal{T}^{\odot n}$  is a factor of  $\mathcal{S}$  <sup>(20)</sup>. Moreover, the restriction of  $\lambda$  to a joining  $\mathcal{T} \vee \mathcal{T}^{\odot n}$  between  $\mathcal{T}$  and  $\mathcal{T}^{\odot n}$  is such that  $\mathcal{T} \vee \mathcal{T}^{\odot n}$

---

<sup>(20)</sup> By  $\mathcal{T}^{\odot n}$  we denote the symmetric factor of  $\mathcal{T}^{\times n}$ , i.e. the sub- $\sigma$ -algebra of sets invariant under all permutations of coordinates (see e.g. [20]).

is a factor of  $\mathcal{T}^{\times n}$ . Since the extension  $\mathcal{T}^{\times n} \rightarrow \mathcal{T}^{\odot n}$  has finite fibers, it follows that also the intermediate extension  $\mathcal{T} \vee \mathcal{T}^{\odot n} \rightarrow \mathcal{T}^{\odot n}$  has finite fibers (see the proof of Proposition 4.5). Now, applying Lemma 4.11 to  $\tilde{\mathcal{T}} = \mathcal{T} \vee \mathcal{T}^{\odot n}$ , we get  $\mathcal{R} = \mathcal{S}$ , and  $\rho$  being the joining associated with the factoring map, we find that the fibers in the extension  $\mathcal{T} \vee \mathcal{S} \rightarrow \mathcal{S}$  are finite as well. ■

REMARK 4.12. Notice that Proposition 4.3 also follows by combining Propositions 4.5, 4.7 and 4.9. In the extended version of the paper (available on arXiv) we give yet another proof of Proposition 4.3.

**4.2. Non-uniqueness.** Since the R-property implies the 2-QS property, one could ask if a result analogous to Corollary 4.8 also holds for 2-QS flows. It turns out that this is not the case. In fact, we “lose” the uniqueness of the embedding of the time-one map already for 2-fold simple flows.

Let  $D_6$  stand for the smallest non-abelian group, the dihedral group of order 6. Its Cayley table is

$$(4.1) \quad \begin{array}{c|cccccc} * & \mathbf{e} & \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{d} & \mathbf{f} \\ \hline \mathbf{e} & e & a & b & c & d & f \\ \mathbf{a} & a & e & d & f & b & c \\ \mathbf{b} & b & f & e & d & c & a \\ \mathbf{c} & c & d & f & e & a & b \\ \mathbf{d} & d & c & a & b & f & e \\ \mathbf{f} & f & b & c & a & e & d \end{array}$$

PROPOSITION 4.13. *There exists an automorphism which is 2-fold simple and has two non-isomorphic square roots (one 2-fold simple, the other only 2-QS).*

*Proof.* Let  $T$  be a weakly mixing automorphism with the MSJ property and consider  $\varphi: X \rightarrow D_6$  such that  $T_\varphi: X \times D_6 \rightarrow X \times D_6$  is also weakly mixing. Then  $T_\varphi$  is 2-fold simple. For  $g \in D_6$  let  $\sigma_g: X \times D_6 \rightarrow X \times D_6$  be given by  $\sigma_g(x, h) = (x, hg)$ . Notice that for all  $g \in D_6$  we have  $\sigma_g \in C(T_\varphi)$ . From now on we will be using the notation of (4.1) for elements of  $D_6$ . Since  $a^2 = e$ , we have  $(\sigma_a)^2 = \text{Id}$ , whence

$$(4.2) \quad (T_\varphi \circ \sigma_a)^2 = (T_\varphi)^2.$$

Moreover, since  $a * b \neq b * a$ ,  $\sigma_b \notin C(\sigma_a)$ , whence  $\sigma_b \notin C(T_\varphi \circ \sigma_a)$ . Therefore

$$(4.3) \quad C(T_\varphi \circ \sigma_a) \neq C(T_\varphi).$$

By Proposition 3.1, we have

$$(4.4) \quad C(T_\varphi) = C((T_\varphi)^2).$$

Combining (4.2)–(4.4) we conclude that  $C((T_\varphi \circ \sigma_a)^2) \neq C(T_\varphi \circ \sigma_a)$ . It follows by Lemma 3.1 that  $T_\varphi$  is 2-fold simple, whereas  $T_\varphi \circ \sigma_a$  is not 2-fold simple (it is 2-QS by Proposition 3.24). ■

**PROPOSITION 4.14.** *There exists a 2-fold simple automorphism which is embeddable into two non-isomorphic flows (one 2-fold simple and the other one only 2-QS).*

*Proof.* Let  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  be a weakly mixing flow on  $(X, \mathcal{B}, \mu)$  with the MSJ property. Let  $\varphi: X \times \mathbb{R} \rightarrow \mathrm{SU}(2)$  be an  $\mathbb{R}$ -cocycle such that  $\mathcal{T}_\varphi = ((T_t)_\varphi)_{t \in \mathbb{R}}$  is weakly mixing. For  $g \in \mathrm{SU}(2)$  let  $\sigma_g(x, h) = (x, hg)$  and let  $\mathcal{R} = (R_t)_{t \in \mathbb{R}}$  be given by  $R_t = \sigma_{g_t}$ , where

$$g_t = \begin{pmatrix} e^{2\pi i t} & 0 \\ 0 & e^{-2\pi i t} \end{pmatrix}.$$

Let  $\mathcal{T} \circ \mathcal{R} := ((T_t)_\varphi \circ R_t)_{t \in \mathbb{R}}$ . Then  $(T_1)_{\varphi} \circ R_1 = (T_1)_\varphi$ . Moreover, for all  $g \in \mathrm{SU}(2)$  we have  $\sigma_g \in C(\mathcal{T}_\varphi)$ . In particular,  $\sigma_{h_0} \in C(\mathcal{T}_\varphi)$ , where  $h_0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ . On the other hand, easy calculation shows that  $h_0 g_t \neq g_t h_0$  for  $t \notin \frac{1}{2}\mathbb{Z}$ , i.e.  $\sigma_{h_0} \notin C(\mathcal{R})$ , whence  $C(\mathcal{T}_\varphi \circ \mathcal{R}) \neq C(\mathcal{T}_\varphi)$ . Thus, arguing as in the proof of the previous proposition, we obtain  $C(\mathcal{T}_\varphi \circ \mathcal{R}) \neq C((T_1)_\varphi \circ R_1)$  and  $\mathcal{T}_\varphi \circ \mathcal{R}$  is not 2-fold simple (it is 2-QS by Proposition 3.23). ■

**REMARK 4.15.** It follows from the above example that the answer to the following question is negative:

- Suppose that an extension  $\tilde{\mathcal{T}} \rightarrow \mathcal{T}$  is such that  $\tilde{T}_1 \rightarrow T_1$  is a compact group extension. Is it true that also  $\tilde{\mathcal{T}} \rightarrow \mathcal{T}$  is a compact group extension?

Indeed, if the extensions were compact group extensions, then  $T_\varphi \circ \sigma_a$  and  $\mathcal{T} \circ \mathcal{R}$  would be 2-fold simple, which is not the case.

Also the answer to the following question is negative:

- Suppose that the extension  $\tilde{T} \rightarrow T$  is such that  $\tilde{T}^2 \rightarrow T^2$  is a compact group extension. Is it true that also  $\tilde{T} \rightarrow T$  is a compact group extension?

**5. Self-joining properties of  $\mathcal{T}$  and  $T_1$ , part II.** A direct consequence of Proposition 4.13 and 4.14 is that the condition  $C(\mathcal{T}) = C(T_1)$  in Proposition 3.1(ii) or in its counterpart for automorphisms and their roots cannot be omitted. We will now discuss this problem in more detail, providing more counterexamples. Let us however see first that such counterexamples cannot be obtained by taking factors of products of simple systems. In particular, products of horocycle flows, their factors and their weakly mixing distal extensions cannot be used.

PROPOSITION 5.1. *Let  $\mathcal{T}$  be a 2-fold simple flow with  $T_1$  ergodic. Then any factor  $\mathcal{A}$  has the following property:  $\mathcal{T}|_{\mathcal{A}}$  and  $T_1|_{\mathcal{A}}$  are either both 2-fold simple, or both 2-QS but not 2-fold simple.*

*Proof.* Notice that  $\mathcal{T}$  and  $T_1$  have the same factors. Moreover, every factor corresponds to a compact subgroup of  $C(\mathcal{T}) = C(T_1)$ . If this subgroup is normal, then  $\mathcal{T}$  and  $T_1$  restricted to the corresponding factor are 2-fold simple. Otherwise they are both only 2-QS but not 2-fold simple. ■

COROLLARY 5.2. *Let  $\mathcal{T}^{(i)}$  be 2-fold simple, weakly mixing flows. Let  $\mathcal{A}$  be a factor of  $\mathcal{T}^{(1)} \times \mathcal{T}^{(2)} \times \dots$  which is 2-QS. Then  $(\mathcal{T}^{(1)} \times \mathcal{T}^{(2)} \times \dots)|_{\mathcal{A}}$  and  $(T_1^{(1)} \times T_1^{(2)} \times \dots)|_{\mathcal{A}}$  are either both 2-fold simple, or both 2-QS but not 2-fold simple.*

*Proof.* This follows immediately from Proposition 5.1 and from the fact that whenever a 2-fold QS system is a factor of a product of two weakly mixing systems, then it is a factor of one of the coordinate factors [19]. ■

Consider the following property of  $\mathcal{T}$ :

(P)  $T_1$  is simple and  $\mathcal{T}$  is only 2-QS.

COROLLARY 5.3. *Let  $\mathcal{T}$  be a weakly mixing flow with the JP property, satisfying the following condition: for any factor  $\mathcal{A}$  of  $\mathcal{T}$  there exists a factor  $\mathcal{C} \subset \mathcal{A}$  such that  $\mathcal{T}|_{\mathcal{C}}$  has property (P). Then  $\mathcal{T}$  is disjoint from weakly mixing extensions of simple flows.*

*Proof.* Recall from [45] that a JP system is disjoint from a weakly mixing simple system if and only if they do not have a common non-trivial factor. Now it suffices to apply Corollary 5.2 and a result from [26] on lifting disjointness. ■

REMARK 5.4. The assumptions of the above corollary are satisfied when  $\mathcal{T}$  satisfies property (P) and has no non-trivial factors.

**5.1. Easy counterexample.** Del Junco and Rudolph [20] recall an example by Glasner [29] of a weakly mixing group extension of a 2-fold simple  $\mathbb{Z}$ -action which is itself not 2-fold simple. It is of the form  $R(x, z_1, z_2) = (Tx, \varphi(x)z_1, z_1z_2)$ , where  $T$  is an ergodic automorphism with minimal self-joinings and  $\varphi: X \rightarrow \mathbb{T}$  yields a weakly mixing group extension  $T_\varphi$  of  $T$ . An example of an ergodic self-joining which is “2 : 1” is given also in [20], showing that this automorphism is not 2-fold simple. We will follow a similar scheme and give an example of a  $\mathbb{Z}_2$ -extension <sup>(21)</sup> of a 2-fold simple map such that the resulting automorphism is “2 : 1” and its square is 2-fold simple.

Let us first recall some necessary lemmas and tools. Since we will use them for abelian groups only, we state them in that setting.

---

<sup>(21)</sup> We will identify  $\mathbb{Z}_2$  with  $\{0, 1\}$  equipped with addition modulo 2.

LEMMA 5.5 ([20]). *Let  $T_\varphi$  be an ergodic compact group extension of a 2-fold simple automorphism  $T$ . Then  $T_\varphi$  is 2-fold simple if and only if for every  $S \in C(T)$  there exists  $\tilde{S} \in C(T_\varphi)$  which is an extension of  $S$ .*

LEMMA 5.6 ( $\mathbb{Z}$ -actions: [54]; abelian groups: [49]; general case: [72]). *Suppose that  $T$  is 2-fold simple and let  $T_\varphi$  be its compact group extension, i.e.  $T_\varphi: X \times G \rightarrow X \times G$ ,  $T_\varphi(x, g) = (Tx, \varphi(x) + g)$ . Let  $\tilde{S} \in C(T_\varphi)$ . Then there are a continuous group automorphism  $v: G \rightarrow G$ , a measurable map  $f: X \rightarrow G$  and  $S \in C(T)$  such that  $\tilde{S}(x, g) = (Sx, f(x) + v(g))$ . Equivalently,  $\varphi(x) - v(\varphi(x)) = f(Tx) - f(x)$ .*

REMARK 5.7. If  $G = \mathbb{Z}_2$  then  $v(i) = i$  and  $f = \text{const}$ , i.e. any  $\tilde{S} \in C(T_\varphi)$  is of the form  $\tilde{S}(x, g) = (Sx, g)$  or  $\tilde{S}(x, g) = (Sx, g + 1)$  for some  $S \in C(T)$ .

LEMMA 5.8 (see e.g. [36]). *Let  $T_\varphi$  be an ergodic compact group extension of  $T$ . Then  $c \in \mathbb{T}$  is an eigenvalue of  $T_\varphi$  if and only if there exist a measurable function  $f: X \rightarrow \mathbb{T}$  and a character  $\chi \in \hat{G}$  such that  $\chi \circ \varphi = c \cdot f \circ T / f$ .*

Let  $T: (X, \mathcal{B}, \mu) \rightarrow (X, \mathcal{B}, \mu)$  be weakly mixing and have the MSJ property, and let  $\varphi: X \rightarrow \mathbb{Z}_2$  yield a weakly mixing group extension  $T_\varphi$ . Consider  $\psi: X \times \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$  given by the formula  $\psi(x, s) = s$ . Let

$$\bar{T}(x, s, r) := (T_\varphi)_\psi(x, s, r) = (Tx, \varphi(x) + s, s + r).$$

LEMMA 5.9.  $\bar{T}$  is weakly mixing.

*Proof.* Suppose that  $\bar{T}$  is not weakly mixing. By Lemma 5.8 there exist  $c \in \mathbb{T} \setminus \{1\}$ ,  $\xi: X \times \mathbb{Z}_2 \rightarrow \mathbb{T}$  and  $\chi \in \hat{\mathbb{Z}}_2$  satisfying the equation

$$(5.1) \quad \chi \circ \psi = c \cdot \xi \circ T_\varphi / \xi.$$

There are two possibilities: (a)  $\chi \equiv 1$ , (b)  $\chi(s) = (-1)^s$ .

In case (a) equation (5.1) takes the form  $\xi \circ T_\varphi = \bar{c} \cdot \xi$ , i.e.  $\bar{c}$  is an eigenvalue of  $T_\varphi$ . This however is not possible, as  $T_\varphi$  is weakly mixing.

In case (b) equation (5.1) takes the form  $(-1)^s = c \cdot \xi \circ T_\varphi / \xi$ . Hence  $\bar{c}^2 \cdot \xi^2 = \xi^2 \circ T_\varphi$ , which implies  $c^2 = 1$  and  $\xi^2 = 1$  almost everywhere as  $T_\varphi$  is weakly mixing (i.e. we have  $|\xi| = 1$  almost everywhere). Now, either (b1)  $c = 1$ , or (b2)  $c = -1$ . Equation (5.1) yields

$$(-1)^s = \xi \circ T_\varphi / \xi \quad \text{in case (b1)} \quad \text{and} \quad (-1)^{s+1} = \xi \circ T_\varphi / \xi \quad \text{in case (b2)},$$

i.e.

$$\begin{aligned} (-1)^s \xi(x, s) &= \xi \circ T_\varphi(x, s) = \xi(Tx, \varphi(x) + s) && \text{in case (b1),} \\ (-1)^{s+1} \xi(x, s) &= \xi \circ T_\varphi(x, s) = \xi(Tx, \varphi(x) + s) && \text{in case (b2).} \end{aligned}$$

Using the theory of characters ( $\xi$  is a square-integrable function) we get

$$\xi(x, s) = \xi_1(x) + \xi_2(x)(-1)^s$$

for some  $\xi_j: X \rightarrow \mathbb{C}$ ,  $j = 1, 2$ . Therefore

$$(-1)^s(\xi_1(x) + \xi_2(x) \cdot (-1)^s) = \xi_1(Tx) + \xi_2(Tx) \cdot (-1)^{\varphi(x)+s} \quad \text{in case (b1),}$$

$$(-1)^{s+1}(\xi_1(x) + \xi_2(x) \cdot (-1)^s) = \xi_1(Tx) + \xi_2(Tx) \cdot (-1)^{\varphi(x)+s} \quad \text{in case (b2),}$$

whence

$$2\xi_2(x) - \xi_1(Tx) = (-1)^s(-\xi_1(x) + \xi_2(Tx) \cdot (-1)^{\varphi(x)}) \quad \text{in case (b1),}$$

$$-\xi_2(x) - \xi_1(Tx) = (-1)^s(\xi_1(x) + \xi_2(Tx) \cdot (-1)^{\varphi(x)}) \quad \text{in case (b2).}$$

Therefore (the above equations hold for all  $s$ )

$$\xi_2(x) = \begin{cases} \xi_1(Tx) & \text{for almost every } x \text{ in case (b1),} \\ \xi_1(Tx) & \text{for almost every } x \text{ in case (b2).} \end{cases}$$

This implies

$$\xi_1(x) = \begin{cases} \xi_2(Tx) \cdot (-1)^{\varphi(x)} = \xi_1(T^2x) \cdot (-1)^{\varphi(x)} & \text{in case (b1),} \\ -\xi_2(Tx) \cdot (-1)^{\varphi(x)} = \xi_1(T^2x) \cdot (-1)^{\varphi(x)} & \text{in case (b2),} \end{cases}$$

whence  $|\xi_1|$  is constant. In either case we have

$$\frac{\xi_1(x) \cdot \xi_1(Tx)}{\xi_1(Tx) \cdot \xi_1(T^2x)} = \frac{\xi_1(x)}{\xi_1(T^2x)} = (-1)^{\varphi(x)},$$

which means that  $\chi \circ \varphi$  is a coboundary. By Lemma 5.8, this yields a contradiction as  $T_\varphi$  is ergodic. ■

PROPOSITION 5.10.  $\bar{T}$  is not 2-fold simple, whereas  $\bar{T}^2$  is 2-fold simple.

*Proof.* To show that  $\bar{T}$  is not 2-fold simple, by Lemma 5.5, it suffices to find an element of  $C(T_\varphi)$  which cannot be lifted to an element of  $C(\bar{T})$ . We claim that  $\sigma(x, s) = (x, s + 1)$  is such an automorphism. Suppose that we can find  $\bar{\sigma} \in C(\bar{T})$  which projects down to  $\sigma$ . By Lemma 5.6, we can solve for  $F$  the following equation:

$$F \circ T_\varphi - F = \psi \circ \sigma - v \circ \psi,$$

where  $v \in \text{Aut}(\mathbb{Z}_2)$  and  $F: X \times \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$  is measurable. Since the group operations are in  $\mathbb{Z}_2$ , we have

$$\begin{aligned} F \circ T_\varphi(x, s) - F(x, s) &= \psi \circ \sigma(x, s) + \psi(x, s) \\ &= \psi(x, s + 1) + \psi(x, s) = s + 1 + s = 1. \end{aligned}$$

Applying to this formula  $\chi(r) = (-1)^r$ , we obtain

$$\frac{(-1)^{F \circ T_\varphi}}{(-1)^F} = -1,$$

which is not possible as  $T_\varphi$  is weakly mixing.

We will now show that  $\bar{T}^2$  is 2-fold simple. We have  $\bar{T}^2 = ((T_\varphi)^2)_{\psi + \psi \circ T_\varphi}$ . Since  $T_\varphi$  is 2-fold simple, Proposition 3.1 implies that  $(T_\varphi)^2$  is also 2-fold



simple and, by Lemma 5.5, it suffices to show that all elements of  $C((T_\varphi)^2)$  can be lifted to  $C(\overline{T}^2)$ . By Proposition 3.1, we have  $C((T_\varphi)^2) = C(T_\varphi)$ . Therefore (see Remark 5.7) it suffices to show that  $\sigma(x, s) = (x, s + 1)$  lifts to an element of  $C(\overline{T}^2)$ , i.e. we need to find a measurable solution  $F: X \times \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$  for the following cocycle equation:

$$(5.2) \quad \psi^{(2)} \circ \sigma(x, s) + \psi^{(2)}(x, s) = F \circ (T_\varphi)^2(x, s) + F(x, s),$$

where  $\psi^{(2)} = \psi + \psi \circ T_\varphi$ . However,

$$\begin{aligned} \psi^{(2)}(x, s + 1) + \psi^{(2)}(x, s) &= \psi(x, s + 1) + \psi(T_\varphi(x, s + 1)) + \psi(x, s) + \psi(T_\varphi(x, s)) \\ &= s + 1 + \varphi(x) + s + 1 + s + \varphi(x) + s = 0, \end{aligned}$$

so it suffices to take  $F \equiv 0$  to complete the proof. ■

**5.2. Advanced counterexample.** We will construct a 2-QS flow which is disjoint from simple flows and whose time-one map is 2-fold simple.

**5.2.1. (C, F)-constructions.** (C, F)-constructions for amenable, unimodular, locally compact second countable (l.c.s.c.) groups were introduced in [12] (this is the most general setting appearing in the literature). We recall here this type of construction, correcting a small error appearing in [12] <sup>(22)</sup>.

Let  $G$  be a unimodular l.c.s.c. group. Given two subsets  $E, F \subset G$ , by  $EF$  we mean their product:  $EF = \{ef : e \in E, f \in F\}$ . If  $E = \{e\}$ , we write  $eF$  for  $EF$ . The set  $\{e^{-1} : e \in E\}$  is denoted by  $E^{-1}$ . To define a (C, F)-action of  $G$  we need two sequences  $(F_n)_{n \geq 0}$  and  $(C_n)_{n > 0}$  of subsets of  $G$ , satisfying additional conditions:

$$(5.3) \quad (F_n)_{n \geq 0} \text{ is a Følner sequence in } G,$$

$$(5.4) \quad C_n \text{ is finite and } \#C_n > 1,$$

$$(5.5) \quad F_n C_{n+1} \subset F_{n+1},$$

$$(5.6) \quad F_n c \cap F_n c' = \emptyset \text{ for all } c \neq c' \in C_{n+1}.$$

We set  $X_n := F_n \times \prod_{k > n} C_k$ , endow  $X_n$  with the standard Borel product  $\sigma$ -algebra and define a Borel embedding  $X_n \rightarrow X_{n+1}$  by setting

$$(5.7) \quad (f_n, c_{n+1}, c_{n+2}, \dots) \mapsto (f_n c_{n+1}, c_{n+2}, \dots).$$

It is well-defined due to (5.5). Then we have  $X_1 \subset X_2 \subset \dots$ . Hence  $X := \bigcup_n X_n$  endowed with the natural Borel  $\sigma$ -algebra, say  $\mathcal{B}$ , is a standard Borel space. Given a Borel subset  $A \subset F_n$ , we define

$$[A]_n := \{x \in X : x = (f_n, c_{n+1}, c_{n+2}, \dots) \in X_n \text{ and } f_n \in A\}$$

---

<sup>(22)</sup> The function  $K_m \times D_m^{(n)} \times D_m^{(n)} \in (g, x) \mapsto T_{m,g}^{(n)} x \in R_m^{(n)}$  defined in [12] seems to have a wrong target set. We bypass this problem in (5.9) and (5.10).

and call this set an  $n$ -cylinder. It is clear that the  $\sigma$ -algebra  $\mathcal{B}$  is generated by the family of all cylinders.

Now we are going to define a “canonical” measure on  $(X, \mathcal{B})$ . Let  $\kappa_n$  stand for the equidistribution on  $C_n$ , and  $\nu_n := (\#C_1 \dots \#C_n)^{-1} \lambda_G|_{F_n}$  on  $F_n$ . We define a product measure  $\mu_n$  on  $X_n$  by setting  $\mu_n = \nu_n \times \kappa_{n+1} \times \kappa_{n+2} \times \dots$ ,  $n \in \mathbb{N}$ . Then the embeddings (5.7) are measure-preserving. Hence a  $\sigma$ -finite measure  $\mu$  on  $X$  is well-defined by the restrictions  $\mu|_{X_n} = \mu_n$ ,  $n \in \mathbb{N}$ . Since

$$\mu_{n+1}(X_{n+1}) = \frac{\nu_{n+1}(F_{n+1})}{\nu_{n+1}(F_n C_{n+1})} \mu_n(X_n) = \frac{\lambda_G(F_{n+1})}{\lambda_G(F_n) \#C_{n+1}} \mu_n(X_n),$$

it follows that  $\mu$  is finite if and only if

$$(5.8) \quad \prod_{n=0}^{\infty} \frac{\lambda_G(F_{n+1})}{\lambda_G(F_n) \#C_{n+1}} < \infty, \quad \text{i.e.} \quad \sum_{n=0}^{\infty} \frac{\lambda_G(F_{n+1} \setminus (F_n C_{n+1}))}{\lambda_G(F_n) \#C_{n+1}} < \infty.$$

From now on we will assume that (5.8) is satisfied. Moreover, we normalize  $\lambda_G$  in such a way that  $\mu(X) = 1$ .

To construct a  $\mu$ -preserving action of  $G$  on  $(X, \mu)$ , we fix a filtration  $K_1 \subset K_2 \subset \dots$  of  $G = \bigcup_{m \geq 1} K_m$  by compact sets. Additionally assume that

$$(5.9) \quad K_m K_m \subset K_{m+1} \quad \text{for all } m \geq 1.$$

Given  $n, m \in \mathbb{N}$  set

$$D_m^{(n)} := \left( \bigcap_{k \in K_m} (k^{-1} F_n) \cap F_n \right) \times \prod_{l > n} C_l \subset X_n.$$

It is easy to verify that  $D_{m+1}^{(n)} \subset D_m^{(n)} \subset D_m^{(n+1)}$ . We define a Borel mapping

$$(5.10) \quad K_m \times D_{m+1}^{(n)} \ni (g, x) \mapsto T_{m,g}^{(n)} x \in D_m^{(n)}$$

by setting, for  $x = (f_n, c_{n+1}, c_{n+2}, \dots)$ ,

$$T_{m,g}^{(n)}(f_n, c_{n+1}, c_{n+2}, \dots) := (g f_n, c_{n+1}, c_{n+2}, \dots).$$

Indeed, for  $g \in K_m$  and  $f_n \in \bigcap_{k \in K_{m+1}} (k^{-1} F_n) \cap F_n$  we obtain  $g f_n \in F_n$  and

$$g f_n \in \bigcap_{k \in K_{m+1}} g k^{-1} F_n = \bigcap_{k \in K_{m+1} g^{-1}} k^{-1} F_n \subset \bigcap_{k \in K_m} k^{-1} F_n,$$

since  $K_{m+1} g^{-1} \supset K_m$  by (5.9). Now let  $D_m := \bigcup_{n=1}^{\infty} D_m^{(n)}$ . Then a Borel mapping

$$T_{m,g}: K_m \times D_{m+1} \ni (g, x) \mapsto T_{m,g} x \in D_m$$

is well-defined by the restrictions  $T_{m,g}|_{D_m^{(n)}} = T_{m,g}^{(n)}$  for  $g \in K_m$  and  $n \geq 1$ . It is clear that  $D_m \supset D_{m+1}$  and  $T_{m,g}|_{D_{m+2}} = T_{m+1,g}$  for all  $m$ . It follows from (5.3) that  $\mu_n(D_m^{(n)}) \rightarrow 1$  as  $n \rightarrow \infty$ . Hence  $\mu(D_m) = 1$  for all  $m \in \mathbb{N}$ .

Finally we set  $\widehat{X} := \bigcap_{m \geq 1} D_m$  and define a Borel mapping  $T: G \times \widehat{X} \ni (g, x) \mapsto T_g x \in \widehat{X}$  by setting  $T_g x := T_{m,gs}$  for some (and hence any)  $m$  such that  $g \in K_m$ . It is clear that  $\mu(\widehat{X}) = 1$ . Thus,  $\mathfrak{T} := (T_g)_{g \in G}$  is a free Borel measure-preserving action of  $G$  on a conull subset of the standard probability space  $(X, \mathcal{B}, \mu)$ . It is easy to see that  $\mathfrak{T}$  does not depend on the choice of filtration  $(K_m)_{m \geq 1}$ .

DEFINITION 5.1. The action  $\mathfrak{T}$  is called the  $(C, F)$ -action of  $G$  associated with  $(C_n, F_n)_n$ .

We now list some basic properties of  $\mathfrak{T}$ . Given Borel subsets  $A, B \subset F_n$  we have

$$[A \cap B]_n = [A]_n \cap [B]_n, \quad [A \cup B]_n = [A]_n \cup [B]_n,$$

$$[A]_n = [AC_{n+1}]_{n+1} = \bigsqcup_{c \in C_{n+1}} [Ac]_{n+1},$$

$$(5.11) \quad T_g[A]_n = [gA]_n \quad \text{if } gA \subset F_n,$$

$$(5.12) \quad \mu([A]_n) = \#C_{n+1} \cdot \mu([Ac]_{n+1}) \quad \text{for every } c \in C_{n+1},$$

$$(5.13) \quad \mu([A]_n) = \frac{\lambda_G(A)}{\lambda_G(F_n)} \mu(X_n),$$

where  $\sqcup$  stands for disjoint union.

REMARK 5.11.  $(C, F)$ -actions are of so-called *funny rank one* (see [23] for  $\mathbb{Z}$ -actions and [73] for the general case). In particular, they are ergodic.

**5.2.2. Construction of a counterexample.** In this section we will provide an example of a weakly mixing flow  $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$  with  $T_1$  2-fold simple such that  $\mathcal{T}$  is only 2-QS (i.e. not 2-fold simple). Moreover, we will show later that our flow has a factor with the same self-joining properties, which is additionally disjoint from all 2-fold simple flows.

Let  $SU(2)$  stand for the special unitary group of order 2, i.e. the group of complex matrices of the form  $\begin{pmatrix} z & -\bar{w} \\ w & \bar{z} \end{pmatrix}$  with determinant 1. Let  $\lambda_{SU(2)}$  denote the normalized Haar measure on  $SU(2)$ . Let  $\varphi: \mathbb{R} \rightarrow \text{Aut}(SU(2), \lambda_{SU(2)})$  be given by

$$\varphi_t(N) = \begin{pmatrix} e^{\pi it/2} & 0 \\ 0 & e^{-\pi it/2} \end{pmatrix} N \begin{pmatrix} e^{-\pi it/2} & 0 \\ 0 & e^{\pi it/2} \end{pmatrix}.$$

We consider the semidirect product  $G = \mathbb{R} \rtimes_{\varphi} SU(2)$ . Recall that multiplication in  $G$  is given by

$$(t, M)(s, N) = (t + s, M\varphi_t(N)).$$

Notice that

$$C(G) = \{(2n, I), (2n, -I) : n \in \mathbb{Z}\}.$$

Given  $a > 0$  we let

$$I^{\mathbb{Z}}[a] := \{m \in \mathbb{Z} : |m| < a\}, \quad I_+^{\mathbb{Z}}[a] := I^{\mathbb{Z}}[a] \cup \{a\}, \quad I^{\mathbb{R}}[a] := (-a, a).$$

Let  $(r_n)_{n \geq 0}$  be an increasing sequence of positive integers such that

$$(5.14) \quad \lim_{n \rightarrow \infty} n^4/r_n = 0.$$

We will impose later one more restriction on the growth of  $(r_n)_{n \geq 0}$ . We define recursively two other sequences  $(a_n)_{n \geq 0}$  and  $(\tilde{a}_n)_{n \geq 0}$  by setting

$$(5.15) \quad a_0 := \tilde{a}_0 := 1, \quad a_{n+1} := \tilde{a}_n(2r_n - 1), \quad \tilde{a}_{n+1} := a_{n+1} + (2n+1)\tilde{a}_n.$$

For each  $n \in \mathbb{N}$  we let

$$(5.16) \quad \begin{aligned} H &:= \mathbb{Z}, & H_n &:= I^{\mathbb{Z}}[r_n], \\ F_n &:= I^{\mathbb{R}}[a_n] \times \mathrm{SU}(2), & \tilde{F}_n &:= I^{\mathbb{R}}[\tilde{a}_n] \times \mathrm{SU}(2), \\ S_n &:= I^{\mathbb{R}}[(2n-1)\tilde{a}_{n-1}] \times \mathrm{SU}(2). \end{aligned}$$

We also consider the homomorphism  $\phi_n: H \rightarrow \mathbb{Z} \times \mathrm{SU}(2) \subset G$  given by

$$\phi_n(h) = (2h\tilde{a}_n, I).$$

Then we have

$$(5.17) \quad S_n \subset F_n, \quad F_n S_n = S_n F_n \subset \tilde{F}_n \subset G,$$

$$(5.18) \quad F_{n+1} = \bigsqcup_{h \in H_n} \tilde{F}_n \phi_n(h) = \bigsqcup_{h \in H_n} \phi_n(h) \tilde{F}_n,$$

$$(5.19) \quad S_{n+1} = \bigsqcup_{h \in I^{\mathbb{Z}}[n+1]} \tilde{F}_n \phi_n(h) = \bigsqcup_{h \in I^{\mathbb{Z}}[n+1]} \phi_n(h) \tilde{F}_n.$$

Fix a sequence  $\varepsilon_n \rightarrow 0$  as  $n \rightarrow \infty$ .

For any two finite sets  $A, B$  and a map  $\phi: A \rightarrow B$  we define a probability measure on  $B$  by

$$\mathrm{dist}_{a \in A} \phi(a) := \frac{1}{\#A} \sum_{a \in A} \delta_{\phi(a)}.$$

Given two measures  $\kappa, \rho$  on a finite set  $B$  let

$$\|\kappa - \rho\|_1 := \sum_{b \in B} |\kappa(b) - \rho(b)|.$$

For  $n \in \mathbb{N}$  let  $\hat{S}_n$  be as in Proposition A.6 and such that its assertion holds for sets  $A', B'$  which are unions of translations of at most  $\#C_1 \cdot \dots \cdot \#C_{n-1}$  elements from  $\mathcal{A}$  (see the remark after Proposition A.6).

LEMMA 5.12 ([11], see also [17]). *If  $r_n$  is sufficiently large then there exists a map  $s_n: H_n \rightarrow \hat{S}_n$  such that for any  $\delta \geq n^{-2}r_n$ ,*

$$(5.20) \quad \|\mathrm{dist}_{t \in I^{\mathbb{Z}}[\delta]}(s_n(h+t), s_n(h'+t), s_n(h''+t)) - \lambda_{\hat{S}_n} \otimes \lambda_{\hat{S}_n} \otimes \lambda_{\hat{S}_n}\|_1 < \varepsilon_n$$

*whenever  $h, h', h'' \in H_n$  are pairwise distinct and  $\{h, h', h''\} + I^{\mathbb{Z}}[\delta] \subset H_n$ .*

From now on we will assume that  $r_n$  is large enough so that we can use Lemma 5.12. For every  $n \in \mathbb{N}$  we fix a map  $s_n$  whose existence is asserted in the lemma. Without loss of generality we may assume that the following boundary condition holds:

$$s_n(r_n - 1) = s_n(-r_n + 1) = 0.$$

Now we can define a map  $c_{n+1}: H_n \rightarrow F_{n+1}$  by setting  $c_{n+1}(h) := s_n(h)\phi_n(h)$ . We also define  $C_{n+1} := c_{n+1}(H_n)$ .

It is easy to derive from (5.17) and (5.18) that (5.3)–(5.6) are satisfied for the sequence  $(F_n, C_{n+1})_{n \geq 0}$ . Moreover,

$$(5.21) \quad \frac{\lambda(F_{n+1})}{\lambda(F_n) \# C_{n+1}} = \frac{2a_{n+1}}{2a_n(2r_n - 1)} = \frac{\tilde{a}_n}{a_n} = 1 + \frac{(2n-1)\tilde{a}_{n-1}}{a_n} = 1 + \frac{2n-1}{2r_{n-1}-1}.$$

From this and from (5.14) we deduce that (5.8) holds. Hence we can consider the associated  $(C, F)$ -action  $\mathfrak{T}$  of  $G$  on  $(X, \mathcal{B}, \mu)$ .

### Auxiliary lemmas

LEMMA 5.13. *Let  $f = f'\phi_{n-1}(h)$ ,  $\hat{f} = \hat{f}'\phi_{n-1}(\hat{h})$  with  $f', \hat{f}' \in \tilde{F}_{n-1}$  and  $h, \hat{h} \in H$ .*

- (i) *Then  $\tilde{F}_{n-1}\phi_{n-1}(h + I^{\mathbb{Z}}[n-1]) \subset fS_n \cup \hat{f}S_n \subset \tilde{F}_{n-1}\phi_{n-1}(h + I^{\mathbb{Z}}[n+1])$ , whence*

$$\begin{aligned} \lambda(fS_n \setminus \tilde{F}_{n-1}\phi_{n-1}(h + I^{\mathbb{Z}}[n-1])) &\leq 4\lambda(\tilde{F}_{n-1}), \\ \lambda(\hat{f}S_n \setminus \tilde{F}_{n-1}\phi_{n-1}(h + I^{\mathbb{Z}}[n-1])) &\leq 4\lambda(\tilde{F}_{n-1}), \end{aligned}$$

*and  $\lambda(fS_n \triangle \hat{f}S_n) \leq 4\lambda(\tilde{F}_{n-1})$ .*

- (ii) *If, in addition,  $fS_n \subset F_n$  then, for any subset  $A \subset F_{n-1}$ ,*

$$\frac{\lambda(AC_n \cap fS_n)}{\lambda(S_n)} = \lambda_{F_{n-1}}(A) + \bar{o}_n(1).$$

*Proof.* By (5.19) we have

$$fS_n = f'\phi_{n-1}(h)\tilde{F}_{n-1}\phi_{n-1}(I^{\mathbb{Z}}[n]) = f'\tilde{F}_{n-1}\phi_{n-1}(h + I^{\mathbb{Z}}[n]).$$

Since  $\tilde{F}_{n-1}\tilde{F}_{n-1} \subset \bigsqcup_{|i| \leq 1} \tilde{F}_{n-1}\phi_{n-1}(i)$ , there exists a partition of  $\tilde{F}_{n-1}$  into subsets  $A_i$ ,  $|i| \leq 1$ , such that  $f'A_i \subset \tilde{F}_{n-1}\phi_{n-1}(i)$  for any  $i$ . Therefore

$$(5.22) \quad \begin{aligned} fS_n &= \bigsqcup_{|i| \leq 1} f'A_i\phi_{n-1}(-i)\phi_{n-1}(i + h + I^{\mathbb{Z}}[n]) \\ &\subset \bigsqcup_{|i| \leq 1} f'A_i\phi_{n-1}(-i)\phi_{n-1}(h + I^{\mathbb{Z}}[n+1]) = \tilde{F}_{n-1}\phi_{n-1}(h + I^{\mathbb{Z}}[n+1]) \end{aligned}$$

since  $\bigsqcup_{|i| \leq 1} f' A_i \phi_{n-1}(-i) = \widetilde{F}_{n-1}$ . In a similar way we obtain

$$(5.23) \quad f S_n \supset \widetilde{F}_{n-1} \phi_{n-1}(h + I^{\mathbb{Z}}[n-1]).$$

Clearly, (5.22) and (5.23) remain true if we replace  $f$  with  $\widehat{f}$ . The remaining assertions of (i) are direct consequences of (5.22) and (5.23).

Suppose that  $f S_n \subset F_n$ . Then  $K := h + I^{\mathbb{Z}}[n-1] \subset H_{n-1}$ , and using (i) we obtain

$$\begin{aligned} \frac{\lambda(AC_n \cap f S_n)}{\lambda(S_n)} &= \frac{1}{\lambda(S_n)} (\lambda(AC_n \cap \widetilde{F}_{n-1} \phi_{n-1}(K)) \pm 4\lambda(\widetilde{F}_{n-1})) \\ &= \frac{1}{\lambda(S_n)} \sum_{h' \in H_{n-1}} (\lambda(As_{n-1}(h') \phi_{n-1}(h') \cap \widetilde{F}_{n-1} \phi_{n-1}(K)) \pm 4\lambda(\widetilde{F}_{n-1})) \\ &= \frac{1}{\lambda(S_n)} \sum_{k \in K} (\lambda(As_{n-1}(h')) \pm 4\lambda(\widetilde{F}_{n-1})) = \frac{\#K}{\lambda(S_n)} \lambda(A) \pm 4 \frac{\lambda(\widetilde{F}_{n-1})}{\lambda(S_n)} \\ &= \lambda_{F_{n-1}}(A) \cdot \frac{\lambda(F_{n-1})}{\lambda(\widetilde{F}_{n-1})} \cdot \frac{\lambda(\widetilde{F}_{n-1})}{\lambda(S_n)} \cdot \#K \pm 4 \frac{\lambda(\widetilde{F}_{n-1})}{\lambda(S_n)} \\ &= \lambda_{F_{n-1}}(A) \cdot \frac{a_{n-1}}{\widetilde{a}_{n-1}} \cdot \frac{\widetilde{a}_{n-1} \cdot (2n-3)}{(2n-1) \cdot \widetilde{a}_{n-1}} \pm 4 \frac{\widetilde{a}_{n-1}}{(2n-1)\widetilde{a}_{n-1}} = \lambda_{F_{n-1}}(A) + \bar{o}_n(1). \blacksquare \end{aligned}$$

LEMMA 5.14. *Let  $(G, \lambda)$  be a locally compact unimodular group with Haar measure  $\lambda$ . For any measurable sets  $A, B, F, S \subset G$  of finite measure  $\lambda$  we have*

$$\int_{S \times S} \lambda(Av \cap Bw \cap F) dv dw = \int_{A \times B} \lambda(aS \cap bS \cap F) da db.$$

*Proof.* Elementary calculation shows that for any measurable subsets  $C, D$  in  $G$  of finite measure we have  $\int_S \lambda(Cv \cap D) dv = \int_C \lambda(xS \cap D) dx$ . Applying this recursively to  $\int_{S \times S} \lambda(Av \cap Bw \cap F) dv dw$  (first for  $C = A$ ,  $D = Bw \cap F$  and then for  $C = B$ ,  $v = w$  and  $D = aS \cap F$ ), we obtain the desired formula. ■

### 5.2.3. Weak mixing

PROPOSITION 5.15.  $T_{(2,I)}$  (and hence also  $T_{(1,I)}$ ) is weakly mixing.

*Proof.* Let  $g_n = \phi_n(1) = (2\widetilde{a}_n, I) = (2, I)^{\widetilde{a}_n}$ . We will show that  $(g_n)_{n \in \mathbb{N}}$  is mixing for  $T$ , i.e.

$$(5.24) \quad \lim_{n \rightarrow \infty} \mu(T_{g_n} D \cap D') = \mu(D)\mu(D').$$

Since  $(T_g)_{g \in G}$  is of funny rank one, it suffices to prove (5.24) for cylinders  $D, D' \in \mathcal{A}$ . Let  $A = A_{(m)}$ ,  $B = B_{(m)} \subset F_m$  be such that  $A, B \in \mathcal{A}$ . We will

now show that

$$\mu(T_{g_n}[A]_m \cap [B]_m) = \mu([A]_m)\mu([B]_m) + \bar{o}_n(1),$$

where  $\bar{o}_n(1)$  means (here and below) a sequence that goes to 0 as  $n$  goes to infinity and that does not depend on  $A, B$ . For  $n \geq m$  we have  $[A]_m = [A_{(n)}]_n$  and  $[B]_m = [B_{(n)}]_n$  for  $A_{(n)}, B_{(n)} \subset F_n$  such that  $A_{(n+1)} = A_{(n)}C_{n+1}$ ,  $B_{(n+1)} = B_{(n)}C_{n+1}$ . For  $n \geq m$  let

$$\begin{aligned} F'_n &:= F_n \cap \bigcap_{(i_1, M_1), (i_2, M_2) \in S_n} F_n(i_1, M_1)(i_2, M_2)^{-1}, \\ A'_{(n)} &:= A_{(n)} \cap F'_n, \\ B'_{(n)} &:= B_{(n)} \cap F'_n, \\ H'_n &:= H_n \cap (H_n - 1). \end{aligned}$$

Notice that

$$\mu([A_{(n)} \setminus A'_{(n)}]_n) \leq \mu([F_n \setminus F'_n]_n), \quad \mu([B_{(n)} \setminus B'_{(n)}]_n) \leq \mu([F_n \setminus F'_n]_n).$$

We claim that

$$(5.25) \quad \mu([F_n \setminus F'_n]_n) = \bar{o}_n(1).$$

Indeed, notice that  $F'_n = I^{\mathbb{R}}[a_n - 2(2n-1)\tilde{a}_{n-1}] \times \text{SU}(2)$ , whence

$$\begin{aligned} (5.26) \quad \lambda_{F_n}(F'_n) &= \frac{\lambda(F'_n)}{\lambda(F_n)} = \frac{a_n - 2(2n-1)\tilde{a}_{n-1}}{a_n} = 1 - 2\frac{(2n-1)\tilde{a}_{n-1}}{a_n} \\ &= 1 - 2\frac{(2n-1)\tilde{a}_{n-1}}{\tilde{a}_{n-1}(2r_{n-1}-1)} = 1 - 2\frac{2n-1}{2r_{n-1}-1} \rightarrow 1 \end{aligned}$$

by (5.14), and (5.25) follows. Therefore

$$\begin{aligned} (5.27) \quad \mu(T_{g_n}[A_{(n)}]_n \cap [B_{(n)}]_n) &= \mu(T_{g_n}[A'_{(n)}]_n \cap [B'_{(n)}]_n) + \bar{o}_n(1) \\ &= \sum_{h \in H_n} \mu(T_{g_n}[A'_{(n)}]_{c_{n+1}(h)}]_{n+1} \cap [B'_{(n)}]_n) + \bar{o}_n(1). \end{aligned}$$

Notice also that  $g_n \in C(G)$  for all  $n \in \mathbb{N}$ . Hence

$$\begin{aligned} g_n A c_{n+1}(h) &= g_n A s_n(h) \phi_n(h) = A s_n(h) g_n \phi_n(h) = A s_n(h) \phi_n(h+1) \\ &= A s_n(h) s_n(h+1)^{-1} c_{n+1}(h+1) \end{aligned}$$

provided that  $h, h+1 \in H_n$ . Moreover, for all  $h \in H_n$ ,

$$\mu(T_{g_n}[A'_{(n)}]_{c_{n+1}(h)}]_{n+1}) = \frac{\mu([A'_{(n)}]_n)}{\#C_{n+1}} \leq \frac{\mu([F_n]_n)}{\#C_{n+1}} = \frac{\mu(X_n)}{\#C_{n+1}} = \bar{o}_n(1).$$

Therefore

$$\begin{aligned}
 (5.28) \quad & \sum_{h \in H_n} \mu(T_{g_n}[A'_{(n)}c_{n+1}(h)]_{n+1} \cap [B'_{(n)}]_n) \\
 &= \sum_{h \in H'_n} \mu(T_{g_n}[A'_{(n)}c_{n+1}(h)]_{n+1} \cap [B'_{(n)}]_n) \\
 &= \sum_{h \in H'_n} \mu([A'_{(n)}s_n(h)s_n(h+1)^{-1}c_{n+1}(h+1)]_{n+1} \cap [B'_{(n)}]_n).
 \end{aligned}$$

We also have

$$A'_{(n)}s_n(h)s_n(h+1)^{-1}c_{n+1}(h+1) \subset F_n c_{n+1}(h+1) \subset F_{n+1}.$$

Moreover,

$$[B'_{(n)}]_n = \bigsqcup_{\tilde{h} \in H_n} [B'_{(n)}c_{n+1}(\tilde{h})]_{n+1},$$

where  $B'_{(n)}c_{n+1}(\tilde{h}) \subset F_n c_{n+1}(\tilde{h})$  and

$$[A'_{(n)}s_n(h)s_n(h+1)^{-1}c_{n+1}(h+1)]_{n+1} \cap [B'_{(n)}c_{n+1}(\tilde{h})]_{n+1} = \emptyset$$

when  $h+1 \neq \tilde{h}$ . Hence

$$\begin{aligned}
 (5.29) \quad & \sum_{h \in H'_n} \mu([A'_{(n)}s_n(h)s_n(h+1)^{-1}c_{n+1}(h+1)]_{n+1} \cap [B'_{(n)}]_n) \\
 &= \sum_{h \in H'_n} (\mu([A'_{(n)}s_n(h)s_n(h+1)^{-1} \cap B'_{(n)}]_{n+1} c_{n+1}(h+1))_{n+1}) \\
 &= \frac{1}{\#H'_n} \sum_{h \in H'_n} \mu([A'_{(n)}s_n(h)s_n(h+1)^{-1} \cap B'_{(n)}]_n) \\
 &= \frac{1}{\#H'_n} \sum_{h \in H'_n} \lambda_{F_n}(A'_{(n)}s_n(h)s_n(h+1)^{-1} \cap B'_{(n)})\mu(X_n) \\
 &= \frac{1}{\#H'_n} \sum_{h \in H'_n} \lambda_{F_n}(A'_{(n)}s_n(h) \cap B'_{(n)}s_n(h+1)) + \bar{o}_n(1) \\
 &= \frac{1}{\#H'_n} \sum_{h \in H'_n} \lambda_{F_n}(A_{(n)}s_n(h) \cap B_{(n)}s_n(h+1)) + \bar{o}_n(1).
 \end{aligned}$$

Recall that  $A_{(n)} = A_{(n-1)}C_n$ ,  $B_{(n)} = B_{(n-1)}C_n$ , and  $[A_{(n-1)}]_{n-1}$ ,  $[B_{(n-1)}]_{n-1}$  are unions of at most  $\#C_1 \dots \#C_{n-1}$  translations of some elements from  $\mathcal{A}$  (recall that  $[A_{(m)}]_m, [B_{(m)}]_m \in \mathcal{A}$ ). Let  $\xi_n := \text{dist}_{h \in H'_n}(s_n(h), s_n(h+1))$ . It follows from Lemma 5.12 that

$$\|\xi_n - \lambda_{\widehat{S}_n} \otimes \lambda_{\widehat{S}_n}\| < \varepsilon_n = \bar{o}_n(1).$$



Define  $f: S_n \times S_n \rightarrow \mathbb{R}$  by

$$f(v, w) = \lambda_{F_n}(A_{(n)}v \cap B_{(n)}w).$$

Then

$$\begin{aligned} (5.30) \quad & \frac{1}{\#H'_n} \sum_{h \in H'_n} \lambda_{F_n}(A_{(n)}s_n(h) \cap B_{(n)}s_n(h+1)) \\ &= \int_{S_n \times S_n} \lambda_{F_n}(A_{(n)}v \cap B_{(n)}w) d\xi_n(v, w) = \int f d\xi_n = \int f d\lambda_{\widehat{S}_n} \otimes \lambda_{\widehat{S}_n} \pm \varepsilon_n \\ &= \int_{S_n \times S_n} \lambda_{F_n}(A_{(n-1)}C_nv \cap B_{(n-1)}C_nw) d\lambda_{\widehat{S}_n} \otimes \lambda_{\widehat{S}_n}(v, w) \pm \varepsilon_n. \end{aligned}$$

Hence by the choice of  $\widehat{S}_n$  we obtain

$$\begin{aligned} (5.31) \quad & \int_{S_n \times S_n} \lambda_{F_n}(A_{(n-1)}C_nv \cap B_{(n-1)}C_nw) d\lambda_{\widehat{S}_n} \otimes \lambda_{\widehat{S}_n}(v, w) \\ &= \int_{S_n \times S_n} \lambda_{F_n}(A_{(n-1)}C_nv \cap B_{(n-1)}C_nw) d\lambda_{S_n} \otimes \lambda_{S_n}(v, w) \pm \varepsilon_n \\ &= \sum_{h, h' \in H_{n-1}} \int_{S_n \times S_n} \frac{\lambda(A_{(n-1)}c_n(h)v \cap B_{(n-1)}c_n(h')w \cap F_n)}{\lambda(F_n)} \\ & \quad \cdot d\lambda_{S_n}(v) d\lambda_{S_n}(w) \pm \varepsilon_n \\ &= \sum_{h, h' \in H_{n-1}} \int_{A_{(n-1)} \times B_{(n-1)}} \frac{\lambda(ac_n(h)S_n \cap bc_n(h')S_n \cap F_n)}{(\lambda(S_n))^2 \lambda(F_n)} d\lambda(a) d\lambda(b) \pm \varepsilon_n. \end{aligned}$$

For  $a \in A_{(n-1)}$ ,  $b \in B_{(n-1)}$  we have

$$(5.32) \quad ac_n(h) = as_{n-1}(h)\phi_{n-1}(h) \in \widetilde{F}_{n-1}\phi_{n-1}(h),$$

$$(5.33) \quad bc_n(h') = bs_{n-1}(h')\phi_{n-1}(h') \in \widetilde{F}_{n-1}\phi_{n-1}(h').$$

Moreover,

$$(0, I)c_n(h) = c_n(h) = s_{n-1}(h)\phi_{n-1}(h) \in \widetilde{F}_{n-1}\phi_{n-1}(h),$$

$$(0, I)c_n(h') = c_n(h') = s_{n-1}(h')\phi_{n-1}(h') \in \widetilde{F}_{n-1}\phi_{n-1}(h').$$

Therefore by Lemma 5.13(i) we obtain

$$\begin{aligned} (5.34) \quad & \lambda((ac_n(h)S_n \cap bc_n(h')S_n \cap F_n) \Delta (c_n(h)S_n \cap c_n(h')S_n \cap F_n)) \\ & \leq \lambda(ac_n(h)S_n \Delta c_n(h)S_n) + \lambda(bc_n(h')S_n \Delta c_n(h')S_n) \leq 8\lambda(\widetilde{F}_{n-1}). \end{aligned}$$

By (5.32), (5.33) and by Lemma 5.13(i) again,

$$ac_n(h)S_n \cap bc_n(h')S_n \subset \widetilde{F}_{n-1}\phi_{n-1}(h + I^{\mathbb{Z}}[n+1]) \cap \widetilde{F}_{n-1}\phi_{n-1}(h' + I^{\mathbb{Z}}[n+1]).$$

Hence  $ac_n(h)S_n \cap bc_n(h')S_n \neq \emptyset$  only if  $h' - h \in I^{\mathbb{Z}}[2n+1]$ . If the latter is satisfied, we say that  $h$  and  $h'$  are *partners*. Denote by  $P(h)$  the set of

partners of  $h$  from  $H_{n-1}$ . Clearly,  $\#P(h) \leq 4n + 1$ . Therefore we deduce from (5.27)–(5.31) that

$$\begin{aligned}
(5.35) \quad & \mu(T_{g_n}[A]_m \cap [B]_m) \\
&= \sum_{h \in H_{n-1}} \sum_{h' \in P(h)} \frac{\lambda(A_{(n-1)})\lambda(B_{(n-1)})}{(\lambda(S_n))^2\lambda(F_n)} \\
&\quad \cdot (\lambda(c_n(h)S_n \cap c_n(h')S_n \cap F_n) \pm 8\lambda(\tilde{F}_{n-1})) \\
&= \frac{\lambda(A_{(n-1)})\lambda(B_{(n-1)})}{(\lambda(S_n))^2\lambda(F_n)} \sum_{h \in H_{n-1}} \sum_{h' \in P(h)} \lambda(c_n(h)S_n \cap c_n(h')S_n \cap F_n) \\
&\quad \pm \lambda(H_{n-1})(4n + 1) \frac{(\lambda(F_{n-1}))^2}{(\lambda(S_n))^2\lambda(F_n)} \cdot 4\lambda(\tilde{F}_{n-1}) \\
&= \frac{\lambda(A_{(n-1)})\lambda(B_{(n-1)})}{(\lambda(F_{n-1}))^2} \theta_n \pm \lambda(H_{n-1})(4n + 1) \frac{(\lambda(F_{n-1}))^2}{(\lambda(S_n))^2\lambda(F_n)} \cdot 4\lambda(\tilde{F}_{n-1}) \\
&= \lambda_{F_{n-1}}(A_{(n-1)})\lambda_{F_{n-1}}(B_{(n-1)})\theta_n \\
&\quad \pm \lambda(H_{n-1})(4n + 1) \frac{(\lambda(F_{n-1}))^2}{(\lambda(S_n))^2\lambda(F_n)} \cdot 4\lambda(\tilde{F}_{n-1}),
\end{aligned}$$

where  $\theta_n > 0$ . Notice that (5.21) together with (5.14) means that  $a_n/\tilde{a}_n \rightarrow 1$  as  $n \rightarrow \infty$ . Therefore, by (5.15) and (5.16) we have

$$\begin{aligned}
& \lambda(H_{n-1})(4n + 1) \frac{(\lambda(F_{n-1}))^2}{(\lambda(S_n))^2\lambda(F_n)} \cdot 4\lambda(\tilde{F}_{n-1}) \\
&= \frac{(2r_{n-1} - 1)(4n + 1)(2a_{n-1})^2 \cdot 4 \cdot 2\tilde{a}_{n-1}}{((2n - 1)\tilde{a}_{n-1})^2 \cdot 2a_n} \\
&= \frac{16(4n + 1)}{2n - 1} \cdot \left(\frac{a_{n-1}}{\tilde{a}_{n-1}}\right)^2 \cdot \frac{\tilde{a}_{n-1}(2r_{n-1} - 1)}{a_n(2n - 1)} \\
&= \frac{16(4n + 1)}{2n - 1} \cdot \left(\frac{a_{n-1}}{\tilde{a}_{n-1}}\right)^2 \cdot \frac{1}{2n - 1} = \bar{o}_n(1).
\end{aligned}$$

Therefore substituting  $A_{(n-1)} = B_{(n-1)} = F_{n-1}$  in (5.35) (see Remark A.12) and passing to the limit we obtain  $\theta_n \rightarrow 1$  as  $n \rightarrow \infty$ . Therefore, using (5.13), we obtain

$$(5.36) \quad \mu(T_{g_n}[A]_m \cap [B]_m) = \mu([A]_m)\mu([B]_m) + \bar{o}_n(1),$$

and the assertion follows. ■

**5.2.4. Self-joining properties.** Let  $k^* := (1, I)(t, M)(1, I)^{-1} = (t, \varphi_1(M))$ , for  $k = (t, M) \in G$ . Notice that  $(k^*)^* = k$ .

PROPOSITION 5.16. *The transformation  $T_{(1,I)}$  is “2 : 1”. Moreover,  $\mathcal{J}_2^e(T_{(1,I)}) = \{\frac{1}{2}(\mu_{T_k} + \mu_{T_{k^*}}) : k \in G\}$ .*

*Proof.* Take any joining  $\nu \in \mathcal{J}_2^e(T_{(1,I)})$ . Let  $I_n := I^{\mathbb{Z}}[n^{-2}a_n]$ ,  $J_n := I^{\mathbb{Z}}[n^{-2}r_n]$  and  $\Phi_n := I_n + 2\tilde{a}_n J_n$ . We first notice that  $(\Phi_n)_{n=1}^\infty$  is a Følner sequence in  $\mathbb{Z}$ . Since

$$\frac{a_n}{n^2} + \frac{2\tilde{a}_n r_n}{n^2} < \frac{\tilde{a}_n(2r_n + 1)}{n^2} < \frac{2a_{n+1}}{(n+1)^2},$$

it follows that  $\Phi_n \subset I_{n+1} + I_{n+1}$ , whence  $\bigcup_{m=1}^n \Phi_m \subset I_{n+1} + I_{n+1}$ . This implies

$$\#\left(\Phi_{n+1} - \bigcup_{m=1}^n \Phi_m\right) \leq 3\#\Phi_{n+1}$$

for every  $n \in \mathbb{N}$ , i.e. Shulman’s condition [48] is satisfied for  $(\Phi_n)_{n=1}^\infty$ . By [48], the pointwise ergodic theorem holds along  $(\Phi_n)_{n=1}^\infty$  for any ergodic transformation. Since  $T_{(1,I)}$  is ergodic by Proposition 5.15, we have

$$(5.37) \quad \frac{1}{\#\Phi_n} \sum_{i \in \Phi_n} \mathbb{1}_D(T_{(i,I)}x) \mathbb{1}_{D'}(T_{(i,I)}x') \rightarrow \nu(D \times D')$$

as  $n \rightarrow \infty$  for  $\nu$ -a.a.  $(x, x') \in X \times X$  and for all cylinders  $D, D' \subset X$ . We call such  $(x, x')$  a *generic point* for  $(T_{(1,I)} \times T_{(1,I)}, \nu)$ . Fix one of them. Then  $x, x' \in X_n$  for all sufficiently large  $n$  and we have the expansions

$$\begin{aligned} x &= (f_n, c_{n+1}(h_n), c_{n+2}(h_{n+1}), \dots), \\ x' &= (f'_n, c_{n+1}(h'_n), c_{n+2}(h'_{n+1}), \dots) \end{aligned}$$

with  $f_n, f'_n \in F_n$  and  $h_i, h'_i \in H_i$ ,  $i > n$ . We let  $H_n^- := I^{\mathbb{Z}}[(1 - n^{-2})r_n] \subset H_n$ . Then

$$\#H_n^- / \#H_n \geq 1 - n^{-2}.$$

Since the marginals of  $\nu$  are both equal to  $\mu$ , by the Borel–Cantelli lemma we may assume without loss of generality that  $h_n, h'_n \in H_n^-$  for all sufficiently large  $n$ . This implies, in turn, that

$$f_{n+1} = f_n c_{n+1}(h_n) \in \tilde{F}_n \phi_n(H_n^-) \subset I_+^{\mathbb{Z}}[(2r_n(1 - n^{-2}) - 1)\tilde{a}_n] \times \text{SU}(2),$$

and similarly,

$$f'_{n+1} \in I_+^{\mathbb{Z}}[(2r_n(1 - n^{-2}) - 1)\tilde{a}_n] \times \text{SU}(2).$$

Notice that given  $g \in \Phi_n$  we have  $(g, I) = (b, I)\phi_n(t)$  for some uniquely determined  $b \in I_n$  and  $t \in J_n$ . Moreover,  $t + h_n \in H_n$ . We also claim that

$$(5.38) \quad (b, I)f_n S_n S_n = (b, I)f_n S_n S_n^{-1} \subset F_n.$$

To verify this, it suffices to show that

$$\frac{a_n}{n^2} + \left(2r_{n-1} \left(1 - \frac{1}{(n-1)^2}\right) - 1\right) \tilde{a}_{n-1} + 2(2n-1)\tilde{a}_{n-1} < a_n.$$

We will show the following stronger inequality for  $n$  large enough:

$$\frac{a_n}{n^2} + 2r_{n-1} \left( 1 - \frac{1}{(n-1)^2} \right) \tilde{a}_{n-1} + 4n\tilde{a}_{n-1} < a_n,$$

i.e.

$$2r_{n-1} \left( 1 - \frac{1}{(n-1)^2} \right) + 4n < (2r_{n-1} - 1) \left( 1 - \frac{1}{n^2} \right).$$

Indeed, the latter is equivalent to  $1 + 4n - \frac{1}{n^2} < 2r_{n-1} \left( \frac{1}{(n-1)^2} - \frac{1}{n^2} \right)$ , which follows from (5.14) in a routine way. Hence

$$(g, I) f_n s_n(h_n) \phi_n(h_n) = d c_{n+1}(t + h_n),$$

$$(g, I) f'_n s_n(h'_n) \phi_n(h'_n) = d' c_{n+1}(t + h'_n),$$

where  $d = (b, I) f_n s_n(h_n) s_n(t + h_n)^{-1}$ ,  $d' = (b, I) f'_n s_n(h'_n) s_n(t + h'_n)^{-1} \in F_n$  by (5.38).

We consider two cases:

- (i)  $h_n = h'_n$  for all  $n \geq N$ ; then  $f_n f'_n{}^{-1} = f_N f'_N{}^{-1} =: k$  for all  $n \geq N$ ,  
so

$$(b, I) k (b, I)^{-1} = \begin{cases} k & \text{when } b \text{ is even,} \\ k^* & \text{when } b \text{ is odd.} \end{cases}$$

- (ii)  $h_n \neq h'_n$  for infinitely many  $n$ .

Our goal is to prove that

$$\nu = \begin{cases} \frac{1}{2}(\mu_{T_k} + \mu_{T_{k^*}}) & \text{in case (i),} \\ \mu \otimes \mu & \text{in case (ii).} \end{cases}$$

It suffices to prove the equalities of the measures restricted to the cylinder sets from  $\mathcal{A}$ . We may also impose additional assumptions on the cylinders, provided that every set can still be approximated by finite unions of such cylinders.

Take  $[A_{(m)}]_m, [B_{(m)}]_m \in \mathcal{A}$  and for  $n \geq m$  set

$$A_{(n+1)} := A_{(n)} C_{n+1}, \quad B_{(n+1)} := B_{(n)} C_{n+1}.$$

We have

$$\begin{aligned} (5.39) \quad & \frac{1}{\#\Phi_n} \#\{g \in \Phi_n : (T_{(g,I)}x, T_{(g,I)}x') \in [A_{(m)}]_m \times [B_{(m)}]_m\} \\ &= \frac{1}{\#\Phi_n} \#\{g \in \Phi_n : (T_{(g,I)}x, T_{(g,I)}x') \in [A_{(n)}]_n \times [B_{(n)}]_n\} \\ &= \frac{1}{\#I_n} \sum_{b \in I_n} \frac{\#\{t \in J_n : d \in A_{(n)}, d' \in B_{(n)}\}}{\#J_n} \\ &= \frac{1}{\#I_n} \sum_{b \in I_n} \xi_n(A_{(n)}^{-1}(b, I) f_n s_n(h_n) \times B_{(n)}^{-1}(b, I) f'_n s_n(h'_n)), \end{aligned}$$

where  $\xi_n = \text{dist}_{t \in J_n}(s_n(t + h_n), s_n(t + h'_n))$ .

We first consider case (i). Take a cylinder  $D' \in \mathcal{A}$  of order  $m$  which is a finite union of cylinders from  $\mathcal{A}$ . It is clear that  $\tilde{D}' := D' \cap [k^{-1}F_m]_m$  is also such a union. Moreover,

$$\begin{aligned} \mu(D' \triangle \tilde{D}') &= \mu\left(\left[\bigsqcup_{j=1}^{j_0} A_j\right]_m \setminus [k^{-1}F_m]_m\right) \\ &= \frac{\lambda(\bigsqcup_{j=1}^{j_0} A_j \setminus k^{-1}F_m)}{\lambda(F_m)} \mu(X_m) \\ &\leq \frac{\lambda(F_m \setminus k^{-1}F_m)}{\lambda(F_m)} = \mu([F_m \setminus k^{-1}F_m]_m). \end{aligned}$$

Given  $\varepsilon$ , for  $m$  large enough,  $\mu([F_m \setminus k^{-1}F_m]_m) < \varepsilon$  since  $(I^{\mathbb{R}}[a_m])_{m \in \mathbb{R}}$  is a Følner sequence in  $\mathbb{R}$ . Therefore we may additionally assume that every cylinder  $[B]_m \in \mathcal{A}$  satisfies  $kB \subset F_m$  if  $m$  is large enough. Since  $kF_m = k^*F_m$ , it follows immediately that also  $k^*B \subset F_m$ .

It is easy to deduce from Lemma 5.12 that  $\|\xi_n - \Delta\|_1 < \varepsilon_n$  for  $n > N$ , where  $\Delta$  is the probability equidistributed on  $\hat{S}_n \times \hat{S}_n$ . This yields (by the choice of  $\hat{S}_n$ )

$$\begin{aligned} (5.40) \quad &\frac{1}{\#I_n} \sum_{b \in I_n} \xi_n(A_{(n)}^{-1}(b, I) f_n s_n(h_n) \times B_{(n)}^{-1}(b, I) f'_n s_n(h'_n)) \\ &= \frac{1}{\#I_n} \sum_{b \in I_n} \lambda_{\hat{S}_n}(A_{(n)}^{-1}(b, I) f_n s_n(h_n) \cap B_{(n)}^{-1}(b, I) f'_n s_n(h'_n)) \pm \varepsilon_n \\ &= \frac{1}{\#I_n} \sum_{b \in I_n} \lambda_{S_n}(A_{(n)}^{-1}(b, I) f_n s_n(h_n) \cap B_{(n)}^{-1}(b, I) f'_n s_n(h'_n)) \pm 2\varepsilon_n. \end{aligned}$$

We obtain further

$$\begin{aligned} (5.41) \quad &\frac{1}{\#I_n} \sum_{b \in I_n} \lambda_{S_n}(A_{(n)}^{-1}(b, I) f_n s_n(h_n) \cap B_{(n)}^{-1}(b, I) f'_n s_n(h'_n)) \\ &= \frac{1}{\#I_n} \sum_{b \in I_n} \frac{\lambda(A_{(n)} \cap (b, I) f_n f_n'^{-1}(b, I)^{-1} B_{(n)} \cap (b, I) f_n s_n(h_n) S_n)}{\lambda(S_n)} \\ &= \frac{1}{\#I_n} \sum_{2b \in I_n} \frac{\lambda(A_{(n)} \cap kB_{(n)} \cap (2b, I) f_n s_n(h_n) S_n)}{\lambda(S_n)} \\ &\quad + \frac{1}{\#I_n} \sum_{2b+1 \in I_n} \frac{\lambda(A_{(n)} \cap k^*B_{(n)} \cap (2b, I) f_n s_n(h_n) S_n)}{\lambda(S_n)}. \end{aligned}$$

Since  $kB_{(m)} \subset F_m$ , we have  $kB_{(n)} \subset F_n$ . Therefore  $A_{(n-1)}C_n \cap kB_{(n-1)}C_n =$

$(A_{(n-1)} \cap kB_{(n-1)})C_n$ . It follows from Lemma 5.13(ii) and (5.38) that

$$\begin{aligned}
 (5.42) \quad & \frac{\lambda(A_{(n)} \cap kB_{(n)} \cap (2b, I)f_n s_n(h_n)S_n)}{\lambda(S_n)} \\
 &= \frac{\lambda((A_{(n-1)} \cap kB_{(n-1)})C_n \cap (2b, I)f_n s_n(h_n)S_n)}{\lambda(S_n)} \\
 &= \lambda_{F_{n-1}}(A_{(n-1)} \cap kB_{(n-1)}) + \bar{o}_n(1) \\
 &= \frac{\mu([A_{(n-1)} \cap kB_{(n-1)}]_{n-1})}{\mu(X_{n-1})} + \bar{o}_n(1) \\
 &= \mu([A_{(m)}]_m \cap T_k[B_{(m)}]_m) + \bar{o}_n(1).
 \end{aligned}$$

In a similar way we obtain

$$\begin{aligned}
 (5.43) \quad & \frac{\lambda(A_{(n)} \cap k^*B_{(n)} \cap (2b, I)f_n s_n(h_n)S_n)}{\lambda(S_n)} \\
 &= \mu([A_{(m)}]_m \cap T_{k^*}[B_{(m)}]_m) + \bar{o}_n(1).
 \end{aligned}$$

Formulas (5.37) and (5.39)–(5.43) lead to  $\nu = \frac{1}{2}(\mu_{T_k} + \mu_{T_{k^*}})$ .

We now consider case (ii). Lemma 5.12 shows that  $\|\xi_n - \lambda_{\widehat{S}_n} \otimes \lambda_{\widehat{S}_n}\|_1 < \varepsilon_n$  for all such  $n$ . Therefore

$$\begin{aligned}
 (5.44) \quad & \frac{1}{\#I_n} \sum_{b \in I_n} \xi_n(A_{(n)}^{-1}(b, I)f_n s_n(h_n) \times B_{(n)}^{-1}(b, I)f'_n s_n(h'_n)) \\
 &= \frac{1}{\#I_n} \sum_{b \in I_n} \lambda_{\widehat{S}_n}(A_{(n)}^{-1}(b, I)f_n s_n(h_n)) \lambda_{\widehat{S}_n}(B_{(n)}^{-1}(b, I)f'_n s_n(h'_n)) \pm \varepsilon_n \\
 &= \frac{1}{\#I_n} \sum_{b \in I_n} \lambda_{S_n}(A_{(n)}^{-1}(b, I)f_n s_n(h_n)) \lambda_{S_n}(B_{(n)}^{-1}(b, I)f'_n s_n(h'_n)) \pm 2\varepsilon_n,
 \end{aligned}$$

where the last equality follows from our choice of  $\widehat{S}_n$ . Now we derive from Lemma 5.13(ii) that

$$\begin{aligned}
 (5.45) \quad & \lambda_{S_n}(A_{(n)}^{-1}(b, I)f_n s_n(h_n)) = \frac{\lambda(A_{(n)} \cap (b, I)f_n s_n(h_n)S_n)}{\lambda(S_n)} \\
 &= \lambda_{F_{n-1}}(B_{(n-1)}) + \bar{o}_n(1) = \frac{\mu([A_{(n-1)}]_{n-1})}{\mu(X_n)} = \mu([A_{(m)}]_m) + \bar{o}_n(1)
 \end{aligned}$$

and

$$(5.46) \quad \lambda_{S_n}(B_{(n)}^{-1}(b, I)f'_n s_n(h'_n)) = \mu([B_{(m)}]_m) + \bar{o}_n(1).$$

It follows from (5.39)–(5.46) that  $\nu = \mu \otimes \mu$ . ■

PROPOSITION 5.17.  $T_{(1, I)}$  is 3-fold PID.

*Sketch of proof.* Let  $\nu \in \mathcal{J}_3^e(T_{(1,I)})$  be pairwise independent. We will show that  $\nu = \mu^{\otimes 3}$ . To this end, we will follow the same path as in the proof of Proposition 5.16. Fix a generic point  $(x, x', x'')$  for  $(T_{(1,I)}^{\times 3}, \nu)$ . Then  $x, x', x'' \in X_n$  for all sufficiently large  $n$ . We may assume that

$$\begin{aligned} x &= (f_n, c_{n+1}(h_n), c_{n+1}(h_{n+1}), \dots), \\ x' &= (f_n, c_{n+1}(h'_n), c_{n+1}(h'_{n+1}), \dots), \\ x'' &= (f_n, c_{n+1}(h''_n), c_{n+1}(h''_{n+1}), \dots), \end{aligned}$$

with  $f_n, f'_n, f''_n \in F_n$  and  $h_i, h'_i, h''_i \in H_i^-$  for  $i > n$ . Now take

$$[A_{(m)}]_m, [A'_{(m)}]_m, [A''_{(m)}]_m \in \mathcal{A}$$

and set  $A_{(n+1)} = A_{(n)}C_{n+1}$ ,  $A'_{(n+1)} = A'_{(n)}C_{n+1}$ ,  $A''_{(n+1)} = A''_{(n)}C_{n+1}$  for  $n \geq m$ . Then as in the proof of previous proposition we have

$$\begin{aligned} (5.47) \quad & \frac{1}{\#\Phi_n} \#\{g \in \Psi_n : (T_{g,I}x, T_{g,I}x', T_{g,I}x'') \in [A_{(m)}]_m \times [A'_{(m)}]_m \times [A''_{(m)}]_m\} \\ &= \frac{1}{\#I_n} \sum_{b \in I_n} \xi_n(A_{(n)}^{-1}d \times A'_{(n)}^{-1}d \times A''_{(n)}^{-1}d), \end{aligned}$$

where  $d = (b, I)f_n s_n(h_n)$ ,  $d' = (b, I)f'_n s_n(h'_n)$ ,  $d'' = (b, I)f''_n s_n(h''_n)$  and

$$\xi_n = \text{dist}_{t \in J_n}(s_n(t + h_n), s_n(t + h'_n), s_n(t + h''_n)).$$

Next, given  $k > 0$ ,

$$\begin{aligned} & \frac{\mu \otimes \mu(\{(y, y') \in X_k \times X_k : y_i \neq y'_i \text{ for all } i > k\})}{\mu(X_k)\mu(X_k)} \\ &= \prod_{i>k} \frac{\#C_i^2 - \#C_i}{\#C_i} \prod_{i>k} \left(1 - \frac{1}{\#H_i}\right) > 0, \end{aligned}$$

where  $(y_i)_{i \geq k}, (y'_i)_{i \geq k}$  are the ‘‘coordinates’’ of  $y, y' \in X_k = F_k \times C_{k+1} \times C_{k+2} \times \dots$  respectively. Since  $\nu$ -almost every point is generic for  $(T_{(1,I)}^{\times 3}, \nu)$ , we can select  $(x, x', x'')$  in such a way that  $h_i, h'_i, h''_i$  are pairwise distinct for all  $i \geq n$  whenever  $n$  is large enough. Therefore, it follows from Lemma 5.12 that  $\|\xi_i - \lambda_{\widehat{S}_i} \otimes \lambda_{\widehat{S}_i} \otimes \lambda_{\widehat{S}_i}\| < \varepsilon_i$  for all  $i \geq n$ . Now we derive from (5.47) and Lemma 5.13 (bearing in mind the choice of  $\widehat{S}_i$ ) that

$$\nu([A_{(m)}]_m \times [A'_{(m)}]_m \times [A''_{(m)}]_m) = \mu([A_{(m)}]_m)\mu([A'_{(m)}]_m)\mu([A''_{(m)}]_m) + \bar{o}_n(1).$$

This implies  $\nu = \mu^{\otimes 3}$ . ■

REMARKS. (1) In view of Proposition 5.16 and the definition of  $k^*$  (see page 56), one can expect that  $T_{(2,I)}$  is 2-fold simple. This is indeed the case (the proof uses similar arguments to those for Proposition 5.16). Moreover,

since  $T_{(1,I)}$  is 3-fold PID, so is  $T_{(2,I)}$  by [11]. It follows by [28] that  $T_{(2,I)}$  is simple (and that  $T_{(1,I)}$  is also PID).

(2) It follows from Propositions 5.16 and 3.23 that the constructed example has as its subaction the flow  $(T_{(t,I)})_{t \in \mathbb{R}}$  which is only QS (it is not simple) with a simple time-2 map  $T_{(2,I)}$ . However, we cannot say at this stage that the flow is disjoint from simple flows (cf. Corollary 5.3)—it is only clear that it has no simple factors (see the next remark).

(3) Recall that since  $T_{(2,I)}$  is simple, any of its factors is determined by a compact subgroup of  $C(T_{(2,I)}) = G$ . The largest compact subgroup of  $G$  is  $\{0\} \times \text{SU}(2) \simeq \text{SU}(2)$  and it is normal in  $G$ . It determines the smallest non-trivial factor of  $T_1$ ; we will denote it by  $\mathcal{A}(\text{SU}(2)) =: \mathcal{A}$ . Notice that  $\mathcal{A}$  is also a factor of  $\mathcal{T}$ . Take  $g \in G \setminus \{0\} \times \text{SU}(2)$ . Then  $\frac{1}{2}(\mu_{T_g} + \mu_{T_{g^*}})|_{\mathcal{A} \otimes \mathcal{A}}$  is a 2-fold ergodic joining of  $\mathcal{T}|_{\mathcal{A}}$ , and it is not difficult to see that it is “2 : 1” over its marginals (this argument is taken from [11]). Since  $\mathcal{T}|_{\mathcal{A}}$  has no non-trivial factors, in view of Corollary 5.3 we obtain an example of a flow disjoint from simple flows, which is QS and whose time-2 map is simple.

(4) An easy linear time change now yields a flow  $\mathcal{T}$  which is QS, is disjoint from simple flows, and has  $T_1$  simple. In particular,  $T_1$  cannot be embedded into a simple flow.

## Appendix. Uniform distribution

**A.1. Necessary tools.** All the information in this section is taken from [40], unless stated otherwise. We will often use it freely, without precise reference.

DEFINITION A.1. Let  $X$  be a compact Hausdorff space with a (regular) probability Borel measure  $\mu$ . A sequence  $(x_n)_{n \in \mathbb{N}} \subset X$  is said to be *uniformly distributed* in  $(X, \mu)$  if for every continuous function  $f: X \rightarrow \mathbb{R}$  we have  $\lim_{N \rightarrow \infty} (1/N) \sum_{n \leq N} f(x_n) = \int_X f d\mu$ .

DEFINITION A.2. We say that  $(x_n)_{n \in \mathbb{N}} \subset \mathbb{R}^s$  is *uniformly distributed modulo 1* if  $(\{x_n\})_{n \in \mathbb{N}}$  is uniformly distributed in  $[0, 1]^s$  with respect to the Lebesgue measure.

DEFINITION A.3. Let  $(x_n)_{n \in \mathbb{N}} \subset \mathbb{R}^s$ . *Discrepancy* is classically defined in the following two ways: for  $N \in \mathbb{N}$  we let

$$D_N^* = D_N^*(x_1, \dots, x_N) := \sup_{J^*} \left| \frac{1}{N} \sum_{n \leq N} \mathbb{1}_{J^*}(\{x_n\}) - \lambda_{\mathbb{R}^s}(J^*) \right|,$$

$$D_N = D_N(x_1, \dots, x_N) := \sup_J \left| \frac{1}{N} \sum_{n \leq N} \mathbb{1}_J(\{x_n\}) - \lambda_{\mathbb{R}^s}(J) \right|.$$

where  $J^*$  runs over all subsets of  $[0, 1]^s$  of the form  $[0, \beta_1) \times \dots \times [0, \beta_s)$ , and  $J$  runs over all subsets of  $[0, 1]^s$  of the form  $[\alpha_1, \beta_1) \times \dots \times [\alpha_s, \beta_s)$ .



PROPOSITION A.1. *Let  $(x_n)_{n \in \mathbb{N}} \subset \mathbb{R}^s$ . The following are equivalent:*

- $(x_n)_{n \in \mathbb{N}}$  is uniformly distributed modulo 1,
- $D_N \rightarrow 0$  as  $N \rightarrow \infty$ ,
- $D_N^* \rightarrow 0$  as  $N \rightarrow \infty$ .

DEFINITION A.4. Let  $\mathcal{S} = \{(x_{n,\sigma}) \subset X : \sigma \in J\}$  be a family of sequences indexed by  $J$ . Then  $\mathcal{S}$  is said to be *equi- $\mu$ -uniformly distributed* in  $X$  if for every  $f \in C(X)$  we have  $\lim_{N \rightarrow \infty} \sup_{\sigma \in J} |\sum_{n \leq N} f(x_{n,\sigma}) - \int f d\mu| = 0$ .

PROPOSITION A.2 ([40]). *Let  $X$  be a compact Hausdorff uniform space. Suppose that  $\{P_\sigma : \sigma \in J\}$  is a family of measure-preserving transformations on  $X$  (with respect to a given  $\mu \in M_1(X)$ ) that is equicontinuous at every point  $x \in X$ , and suppose that  $(x_n)_{n \in \mathbb{N}} \subset X$  is  $\mu$ -uniformly distributed. Then the family  $\{(P_\sigma x_n) \subset X : \sigma \in J\}$  is equi- $\mu$ -uniformly distributed.*

THEOREM A.3 ([34], see also [40]). *Let  $(x_n)_{n \in \mathbb{N}} \subset \mathbb{R}^s$  be uniformly distributed modulo 1 and let  $f: [0, 1]^s \rightarrow \mathbb{R}$  be continuous. Let*

$$D_N^*(f, (x_n)_{1 \leq n \leq N}) := \left| \frac{1}{N} \sum_{i=1}^N f(x_i) - \int_{[0,1]^s} f d\lambda_{\mathbb{R}^s} \right|.$$

Then

$$D_N^*(f, (x_n)_{1 \leq n \leq N}) \leq (1 + 2^{2s-1})M\left(\left[D_N^*(f, (x_n)_{1 \leq n \leq N})\right]^{-1/k}\right)^{-1},$$

where  $M$  is the modulus of continuity of  $f$ .

COROLLARY A.4. *Let  $(x_n)_{n \in \mathbb{N}} \subset \mathbb{R}^s$  be uniformly distributed modulo 1 and let  $\mathcal{F}$  be a family of equicontinuous functions on  $[0, 1]^s$ . Then*

$$\sup_{f \in \mathcal{F}} D_N^*(f, (x_n)_{n \in \mathbb{N}}) \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

THEOREM A.5 (von Neumann, Oxtoby, Ulam, see [62]). *Let  $\mu_1, \mu_2$  be non-atomic probability Borel measures on  $[0, 1]^s$  with full support and such that  $\mu_i(\partial([0, 1]^s)) = 0$ ,  $i = 1, 2$ . Then there exists a homeomorphism  $h: [0, 1]^s \rightarrow [0, 1]^s$  such that  $\mu_1 \circ h^{-1} = \mu_2$ . Moreover,  $h|_{\partial([0, 1]^s)} = \text{id}$ .*

**A.2. Technical results.** Let  $G = \mathbb{R} \times_{\varphi} \text{SU}(2)$ . By  $\lambda_G, \lambda_{\text{SU}(2)}$  we will denote the Haar measures on  $G$  and  $\text{SU}(2)$  respectively such that  $\lambda_G((0, 1) \times \text{SU}(2)) = 1$  and  $\lambda_{\text{SU}(2)}(\text{SU}(2)) = 1$ . Given a measurable set  $S \subset G$ , we denote by  $\lambda_S$  the conditional probability measure on  $S$ . Finally, given a finite set  $\widehat{S} \subset G$ , we denote by  $\lambda_{\widehat{S}}$  the normalized counting measure, i.e.  $\lambda_{\widehat{S}}(A) = \#(\widehat{S} \cap A) / \#\widehat{S}$ . Finally, let  $\varepsilon_n > 0$  for  $n \in \mathbb{N}$ .

PROPOSITION A.6. *There exists a dense family  $\mathcal{A}$  of subsets of  $G = \mathbb{R} \rtimes_{\varphi} \mathrm{SU}(2)$ , and for any  $n \in \mathbb{N}$  there exists a finite subset  $\widehat{S}_n \subset S_n$  such that:*

(i) *For all  $A, B \in \mathcal{A}$  and all  $a, b \in G$  we have*

$$\lambda_{S_n}(A^{-1}a \cap B^{-1}b) = \lambda_{\widehat{S}_n}(A^{-1}a \cap B^{-1}b) \pm \varepsilon_n.$$

(ii) *For all  $A, B \in \mathcal{A}$  and all  $a, b \in G$  we have*

$$\int_{S_n \times S_n} f d\lambda_{S_n} \otimes \lambda_{S_n} = \int_{S_n \times S_n} f d\lambda_{\widehat{S}_n} \otimes \lambda_{\widehat{S}_n} \pm \varepsilon_n$$

where  $f: G \times G \rightarrow \mathbb{R}$  is given by  $f(v, w) = \lambda_{F_n}(Aav \cap Bbw)$ .

Before we begin the proof, we provide some necessary lemmas. Let  $\mathcal{C}_d$  stand for the family of cubes in  $[0, 1]^d$ . By  $\mathbf{d}: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  we denote the maximum metric in  $\mathbb{R}^d$  and by  $\mathbf{d}(x, C)$  the distance from the point  $x$  to the set  $C$ . Given  $C \in \mathcal{C}_d$  and  $\varepsilon > 0$  let

$$C_\varepsilon := \{x \notin C : \mathbf{d}(x, C) < \varepsilon\} \quad \text{and} \quad f_{C, \varepsilon} := \max(1 - \mathbf{d}(x, C)/\varepsilon, 0).$$

Notice that for each  $\varepsilon > 0$ ,

(A.1)  $\{f_{C, \varepsilon} : C \in \mathcal{C}_d\}$  is an equicontinuous family of functions

and

(A.2)  $\lambda(C_\varepsilon) \leq 2d\varepsilon$  for all  $C \in \mathcal{C}_d$  and  $\varepsilon > 0$ .

LEMMA A.7. *Let  $(X, \mu)$  be a compact metric space with a probability Borel measure and let  $(y_n)_{n \in \mathbb{N}} \subset X$  be uniformly distributed with respect to  $\mu$ . Let  $\mathcal{F}$  be a family of measure-preserving equicontinuous homeomorphisms from  $X$  to  $[0, 1]^d$ . Then*

$$\sup_{\Psi \in \mathcal{F}} D_N^*((\Psi(y_n))_{1 \leq n \leq N}) \rightarrow 0.$$

*Proof.* Let  $\varepsilon > 0$  and choose  $0 < \delta < \varepsilon$  small enough that for every cube  $C_\delta$  of edge length at most  $\delta$  we have  $\Psi^{-1}(C_\delta) \subset C'_\varepsilon$ , where  $C'_\varepsilon$  is some cube of edge length at most  $\varepsilon$ . Let  $K := d \cdot 2^{d-1} \cdot 1/\delta$ .

We have

$$\begin{aligned} \left| \frac{1}{N} \sum_{n \leq N} \mathbb{1}_C \circ \Psi(x_n) - \lambda(C) \right| &\leq \underbrace{\left| \frac{1}{N} \sum_{n \leq N} (\mathbb{1}_C \circ \Psi(x_n) - f_{C, \delta} \circ \Psi(x_n)) \right|}_{I_1} \\ &\quad + \underbrace{\left| \frac{1}{N} \sum_{n \leq N} f_{C, \delta} \circ \Psi(x_n) - \int f_{C, \delta} d\lambda \right|}_{I_2} + \underbrace{\left| \int f_{C, \delta} d\lambda - \lambda(C) \right|}_{I_3}. \end{aligned}$$

Notice that  $I_1 \leq \frac{1}{N} \sum_{n \leq N} \mathbb{1}_{C_\delta} \circ \Psi(x_n) = \frac{1}{N} \sum_{n \leq N} \mathbb{1}_{\Psi^{-1}(C_\delta)}(x_n)$ . By the choice of  $K$  the set  $C_\delta$  is included in a union of (at most)  $K$  cubes of edge length at most  $\delta$ . Therefore, by the choice of  $\delta$ ,  $\Psi^{-1}(C_\delta)$  is included in a union of (at most)  $K$  cubes of edge length at most  $\varepsilon$ . It follows immediately that there exists  $N_0$  such that  $I_1 \leq \varepsilon$  for  $N \geq N_0$ .

By Theorem A.3 we have  $I_2 \rightarrow 0$ . Moreover, taking (A.1) into account, we see that the rate of convergence can be estimated using only  $D_N^*((x_n)_{1 \leq n \leq N})$  (and  $\varepsilon$ ).

Finally,  $I_3 \leq \lambda(C_\delta) < 2d\delta \leq 2d\varepsilon$  and the claim follows. ■

LEMMA A.8. *Let  $(y_n)_{n \in \mathbb{N}} \subset [0, 1]^d$  be uniformly distributed. Let  $\mathcal{F}$  be a family of equicontinuous homeomorphisms of  $[0, 1]^d$  preserving the Lebesgue measure. Then*

$$\sup_{\Psi_1, \Psi_2 \in \mathcal{F}} \sup_{C, D, E \subset \mathcal{C}_d} \left| \frac{1}{N} \sum_{n \leq N} \mathbb{1}_{\Psi_1(C) \cap \Psi_2(D) \cap E}(y_n) - \lambda(\Psi_1(C) \cap \Psi_2(D) \cap E) \right| \rightarrow 0.$$

*Proof.* Choose  $\varepsilon > 0$  and let

$$\begin{aligned} I_1 &:= \lambda(\Psi_1(C) \cap \Psi_2(D) \cap E) - \int f_{C, \varepsilon} \circ \Psi_1^{-1} \cdot f_{D, \varepsilon} \circ \Psi_2^{-1} \cdot f_{E, \varepsilon} d\lambda, \\ I_2 &:= \frac{1}{N} \sum_{n \leq N} \mathbb{1}_{\Psi_1(C) \cap \Psi_2(D) \cap E}(x_n) - \frac{1}{N} \sum_{n \leq N} (f_{C, \varepsilon} \circ \Psi_1^{-1} \cdot f_{D, \varepsilon} \circ \Psi_2^{-1} \cdot f_{E, \varepsilon})(x_n), \\ I_3 &:= \frac{1}{N} \sum_{n \leq N} (f_{C, \varepsilon} \circ \Psi_1^{-1} \cdot f_{D, \varepsilon} \circ \Psi_2^{-1} \cdot f_{E, \varepsilon})(x_n) \\ &\quad - \int f_{C, \varepsilon} \circ \Psi_1^{-1} \cdot f_{D, \varepsilon} \circ \Psi_2^{-1} \cdot f_{E, \varepsilon} d\lambda. \end{aligned}$$

It suffices to show that  $I_1, I_2, I_3 \rightarrow 0$  and the convergence is uniform with respect to  $C, D, E \in \mathcal{C}_d$  and  $\Psi_1, \Psi_2 \in \mathcal{F}$ .

We have

$$\begin{aligned} |I_1| &\leq \left| \int (\mathbb{1}_C \circ \Psi_1^{-1} \cdot \mathbb{1}_D \circ \Psi_2^{-1} \cdot \mathbb{1}_E - \mathbb{1}_C \circ \Psi_1^{-1} \cdot \mathbb{1}_D \circ \Psi_2^{-1} \cdot f_{E, \varepsilon}) d\lambda \right| \\ &\quad + \left| \int (\mathbb{1}_C \circ \Psi_1^{-1} \cdot \mathbb{1}_D \circ \Psi_2^{-1} \cdot f_{E, \varepsilon} - \mathbb{1}_C \circ \Psi_1^{-1} \cdot f_{D, \varepsilon} \circ \Psi_2^{-1} \cdot f_{E, \varepsilon}) d\lambda \right| \\ &\quad + \left| \int (\mathbb{1}_C \circ \Psi_1^{-1} \cdot f_{D, \varepsilon} \circ \Psi_2^{-1} \cdot f_{E, \varepsilon} - f_{C, \varepsilon} \circ \Psi_1^{-1} \cdot f_{D, \varepsilon} \circ \Psi_2^{-1} \cdot f_{E, \varepsilon}) d\lambda \right| \\ &\leq \int |\mathbb{1}_E - f_{E, \varepsilon}| d\lambda + \int |\mathbb{1}_D \circ \Psi_2^{-1} - f_{D, \varepsilon} \circ \Psi_2^{-1}| d\lambda \\ &\quad + \int |\mathbb{1}_C \circ \Psi_1^{-1} - f_{C, \varepsilon} \circ \Psi_1^{-1}| d\lambda \\ &= \int |\mathbb{1}_E - f_{E, \varepsilon}| d\lambda + \int |\mathbb{1}_D - f_{D, \varepsilon}| d\lambda + \int |\mathbb{1}_C - f_{C, \varepsilon}| d\lambda \\ &\leq \lambda(E_\varepsilon) + \lambda(D_\varepsilon) + \lambda(C_\varepsilon) \leq 6d\varepsilon, \end{aligned}$$

where the last inequality follows from (A.2).

We will now estimate  $I_2$ :

$$\begin{aligned}
|I_2| &\leq \left| \frac{1}{N} \sum_{n \leq N} [(\mathbb{1}_{\Psi_1(C)} \cdot \mathbb{1}_{\Psi_2(D)} \cdot \mathbb{1}_E)(x_n) - (\mathbb{1}_{\Psi_1(C)} \cdot \mathbb{1}_{\Psi_2(D)} \cdot f_{E,\varepsilon})(x_n)] \right| \\
&\quad + \left| \frac{1}{N} \sum_{n \leq N} [(\mathbb{1}_{\Psi_1(C)} \cdot \mathbb{1}_{\Psi_2(D)} \cdot f_{E,\varepsilon})(x_n) - (\mathbb{1}_{\Psi_1(C)} \cdot f_{D,\varepsilon} \circ \Psi_2^{-1} \cdot f_{E,\varepsilon})(x_n)] \right| \\
&\quad + \left| \frac{1}{N} \sum_{n \leq N} [(\mathbb{1}_{\Psi_1(C)} \cdot f_{D,\varepsilon} \circ \Psi_2^{-1} \cdot f_{E,\varepsilon})(x_n) \right. \\
&\quad \quad \quad \left. - (f_{C,\varepsilon} \circ \Psi_1^{-1} \cdot f_{D,\varepsilon} \circ \Psi_2^{-1} \cdot f_{E,\varepsilon})(x_n)] \right| \\
&\leq \frac{1}{N} \sum_{n \leq N} \mathbb{1}_{E_\varepsilon}(x_n) + \frac{1}{N} \sum_{n \leq N} \mathbb{1}_{\Psi_2(D_\varepsilon)}(x_n) + \frac{1}{N} \sum_{n \leq N} \mathbb{1}_{\Psi_1(C_\varepsilon)}(x_n) \\
&= \frac{1}{N} \sum_{n \leq N} \mathbb{1}_{E_\varepsilon}(x_n) + \frac{1}{N} \sum_{n \leq N} \mathbb{1}_{D_\varepsilon}(\Psi_2^{-1}(x_n)) + \frac{1}{N} \sum_{n \leq N} \mathbb{1}_{C_\varepsilon}(\Psi_1^{-1}(x_n)).
\end{aligned}$$

Since each of the sets  $E_\varepsilon, D_\varepsilon, C_\varepsilon$  is a difference of two cubes, the above expression converges to  $\lambda(E_\varepsilon) + \lambda(D_\varepsilon) + \lambda(C_\varepsilon) \leq 6d\varepsilon$ . Moreover, by Lemma A.7, the convergence is uniform with respect to  $C, D, E \in \mathcal{C}_d$  and  $\Psi_1, \Psi_2 \in \mathcal{F}$ .

$I_3$  also converges to zero and again, by Theorem A.3, the convergence is uniform with respect to  $C, D, E \in \mathcal{C}_d$  and  $\Psi_1, \Psi_2 \in \mathcal{F}$ . ■

Let  $d$  be a right-invariant metric on  $G$ . Let  $\mathcal{U} = \{U_i : 1 \leq i \leq M\}$  be a finite cover of  $\mathrm{SU}(2)$  by open sets such that  $\lambda_{\mathrm{SU}(2)}(\partial U_i) = 0$  for  $1 \leq i \leq M$ . Then

$$(A.3) \quad \mathcal{V} := \bigcup_{l \in \mathbb{Z}} \bigcup_{i=1}^M (l-1, l+2) \times U_i$$

is an open cover of  $G$  such that  $\lambda_G(\partial((l-1, l+2) \times U_i)) = 0$ .

LEMMA A.9. *The cover (A.3) has a positive Lebesgue number.*

*Proof.* Suppose that there is no  $\delta > 0$  such that for each  $(t, \mathbf{M}) \in G$  we have  $B((t, \mathbf{M}), \delta) \subset (l-1, l+2) \times U_i$  for some  $l \in \mathbb{Z}$  and  $1 \leq i \leq M$ . We claim that

$$(A.4) \quad B(x, \varepsilon) \subset (l-1, l+2) \times U_i \Leftrightarrow B(x(-l, I), \varepsilon) \subset (-1, 2) \times U_i.$$

Indeed, assume that  $B((t, \mathbf{M}), \varepsilon) \subset (l-1, l+2) \times U_i$  and take  $y$  in  $B((t, \mathbf{M})(-l, I), \varepsilon)$ . Then  $d(y(l, I), (t, \mathbf{M})) = d(y, (t, \mathbf{M})(-l, I)) < \varepsilon$ , whence

$$y \in [(l-1, l+2) \times U_i](-l, I) = (-1, 2) \times U_i,$$

and (A.4) follows.

According to our assumption there exists  $x_n = (t_n, \mathbf{M}_n)$  such that

(A.5) none of the balls  $B(x_n, 1/n)$  is a subset of an element of our cover.

In view of (A.4), we may assume that  $x_n \in (-1, 2) \times U_i$  for some  $i$ , and also that  $x_n \rightarrow x$  for some  $x = (t, \mathbf{M}) \in (l-1, l+2) \times U_j$ , where  $l \in \mathbb{Z}$ ,  $1 \leq j \leq M$ . Then for  $N$  large enough we have  $B(x, 1/N) \subset (l-1, l+2) \times \bar{U}_j$ . Let  $n \geq 2N$  be large enough that  $x_n \in B(x, 1/(2N))$ . Then  $B(x_n, 1/n) \subset B(x_n, 1/(2N)) \subset B(x, 1/N) \subset (l-1, l+2) \times U_j$ , which contradicts with (A.5). ■

For  $1 \leq i \leq M$  let  $\psi_i: \bar{U}_i \rightarrow [0, 1]^3$  be a homeomorphism which carries  $\lambda_{\bar{U}_i}$  to the Lebesgue measure  $\lambda_{[0,1]^3}$  on  $[0, 1]^3$  <sup>(23)</sup>. For  $1 \leq i \leq M$  and  $l \in \mathbb{Z}$  let  $\psi_{l,i}: [l-1, l+2] \times \bar{U}_i \rightarrow [-1, 2] \times [0, 1]^3$  be

$$\psi_{l,i}(x, \mathbf{M}) = (x-l, \psi_i(\mathbf{M}))$$

and let  $\psi: [-1, 2] \times [0, 1]^3 \rightarrow [0, 1]^4$  be defined by

$$\psi(x_1, x_2, x_3, x_4) = ((x_1+1)/3, x_2, x_3, x_4).$$

For  $a \in G$  define  $g_a, g: G \rightarrow G$  be given by

$$g_a(b) = ba \quad \text{and} \quad g(b) = b^{-1} \quad \text{for } b \in G.$$

LEMMA A.10. *The family of functions*

$$\{g_a \circ g \circ \psi_{l,i}^{-1} \circ \psi^{-1} : a \in G, l \in \mathbb{Z}, 1 \leq i \leq M\}$$

*is uniformly bi-equicontinuous.*

*Proof.* It suffices to show that the following families of functions are uniformly bi-equicontinuous:

- (i)  $\mathcal{F}_1 = \{\psi_{l,i}^{-1} : l \in \mathbb{Z}, 1 \leq i \leq M\}$ ,
- (ii)  $\mathcal{F}_2 = \{g|_{[l-1, l+2] \times \text{SU}(2)} : l \in \mathbb{Z}\}$ ,
- (iii)  $\mathcal{F}_3 = \{g_a : a \in G\}$ .

Notice that  $\mathcal{F}_1$  is a finite family of homeomorphisms of compact spaces, and elements of  $\mathcal{F}_3$  are isometries, so both  $\mathcal{F}_1, \mathcal{F}_3$  are uniformly bi-equicontinuous. We now consider  $\mathcal{F}_2$ . For  $l \in \mathbb{Z}$  let  $g_l = g|_{[l-1, l+2] \times \text{SU}(2)}$ . For  $(t, M), (s, N) \in [l-1, l+2] \times \text{SU}(2)$  we have

$$\begin{aligned} d((t-l, M), (s-l, N)) \\ = d((t, M)(-l, I), (s, N)(-l, I)) = d((t, M), (s, N)) \end{aligned}$$

---

<sup>(23)</sup> Such a homeomorphism exists in view of Theorem A.5.

and

$$\begin{aligned}
d(g_l(t, M), g_l(s, N)) &= d((-t, \varphi_{-t}(M^{-1})), (-s, \varphi_{-s}(N^{-1}))) \\
&= d((-t, \varphi_{-t}(M^{-1}))(l, I), (-s, \varphi_{-s}(N^{-1}))(l, I)) \\
&= d((l-t, \varphi_{-t}(M^{-1})), (l-s, \varphi_{-s}(N^{-1}))) \\
&= d((l-t, \varphi_{l-t}(M^{-1})), (l-s, \varphi_{l-s}(N^{-1}))) \\
&= d(g_0(t-l, M), g_0(s-l, N)).
\end{aligned}$$

Moreover,  $g_l^{-1} = g_{-1-l}$ . ■

Let  $\{\rho_i\}_{1 \leq i \leq M}$  and  $\{\eta_l\}_{l \in \mathbb{Z}}$  be smooth partitions of unity, subordinate to cover  $\mathcal{U}$  and  $\bigcup_{l \in \mathbb{Z}}(l-1, l+2) = \mathbb{R}$  respectively. Clearly,  $\{\eta_l \otimes \rho_i\}_{1 \leq i \leq M, l \in \mathbb{Z}}$  is then a smooth partition of unity subordinate to cover  $\mathcal{V}$ .

Let  $\tilde{\mathcal{A}}$  be a family of “cubes” in  $[0, 1]^4$  sufficiently small so that, for  $a_1, a_2, a_3 \in G$ ,  $l_1, l_2, l_3 \in \mathbb{Z}$ ,  $1 \leq i_1, i_2, i_3 \leq M$ ,  $C_1, C_2, D_1, D_2, D_3 \in \tilde{\mathcal{A}}$  whenever the sets

$$\bigcup_{j=1}^2 g_{a_j} \circ g \circ \psi_{l_j, i_j}^{-1} \circ \psi^{-1}(C_j) \quad \text{and} \quad \bigcup_{j=1}^3 g_{a_j} \circ \psi_{l_j, i_j}^{-1} \circ \psi^{-1}(D_j)$$

are connected then

$$\bigcup_{j=1}^2 g_{a_j} \circ g \circ \psi_{l_j, i_j}^{-1} \circ \psi^{-1}(C_j) \subset (l_C - 1, l_C + 2) \times U_{i_C}$$

and

$$\text{(A.6)} \quad \bigcup_{j=1}^3 g_{a_j} \circ \psi_{l_j, i_j}^{-1} \circ \psi^{-1}(D_j) \subset (l_D - 1, l_D + 2) \times U_{i_D}$$

for some  $l_C, l_D \in \mathbb{Z}$ ,  $1 \leq i_C, i_D \leq M$ . This can be done in view of Lemmas A.9 and A.10.

REMARK A.11. We may assume that in fact a stronger condition than (A.6) holds:

$$\begin{aligned}
\text{(A.7)} \quad & \left( \bigcup_{j=1}^3 g_{a_j} \circ \psi_{l_j, i_j}^{-1} \circ \psi^{-1}(D_j) \right)_{\delta_0} \cup \left( \bigcup_{j=1}^3 g_{a_j} \circ \psi_{l_j, i_j}^{-1} \circ \psi^{-1}(D_j) \right) \\
& \subset (l_D - 1, l_D + 2) \times U_{i_D},
\end{aligned}$$

where  $\delta_0 > 0$  is a sufficiently small number, independent of the choice of  $D_j, a_j$ ,  $1 \leq j \leq 3$ . This can be done for example by adding one more step before fixing the open cover of  $\text{SU}(2)$ . Namely, given  $\{U_i\}_{1 \leq i \leq M}$ , we need to find  $\delta_0 > 0$  such that  $\{(U_i)_{\delta_0} \cup U_i\}_{1 \leq i \leq M}$  still consists of sets homeomorphic to  $(0, 1)^4$  and work with both covers simultaneously.

Let

$$\mathcal{A} = \{\psi_{l,i}^{-1} \circ \psi^{-1}(C) : C \in \tilde{\mathcal{A}}, l \in \mathbb{Z}, 1 \leq i \leq M\}.$$

Let  $(x_n)_{n \in \mathbb{N}} \subset [0, 1] \times \text{SU}(2)$  be uniformly distributed in  $[0, 1] \times \text{SU}(2)$  <sup>(24)</sup>. For  $l \in \mathbb{Z}$  we denote by  $(x_{n,l})_{n \in \mathbb{N}}$  the sequence in  $[l, l+1] \times \text{SU}(2)$  given by  $x_{n,l} = x_n + (l, 0)$  (the addition is understood coordinatewise).

*Proof of Proposition A.6(i).* Fix  $n \in \mathbb{N}$ . The set  $\widehat{S}_n$  we are looking for will consist of elements of sequences  $(x_{n,l})_{n \in \mathbb{N}}$ ; more precisely, we will have

$$\widehat{S}_n = \{x_{n,l} : 1 \leq n \leq n_0, l \in \mathbb{Z}\} \cap S_n,$$

with some  $n_0 \in \mathbb{N}$  sufficiently large (depending on  $\varepsilon_n$ ). Take  $A, B \in \mathcal{A}$  and all  $a, b \in G$ . Then

$$A = \psi_{l_A, i_A}^{-1} \circ \psi^{-1}(C_A), \quad B = \psi_{l_B, i_B}^{-1} \circ \psi^{-1}(C_B)$$

for some  $C_A, C_B \in \tilde{\mathcal{A}}$  and  $l_A, l_B \in \mathbb{Z}$ ,  $1 \leq i_A, i_B \leq M$ .

If  $A^{-1}a \cap B^{-1}b \cap S_n = \emptyset$  then clearly

$$\lambda_{S_n}(A^{-1}a \cap B^{-1}b) = \lambda_{\widehat{S}_n}(A^{-1}a \cap B^{-1}b) = 0,$$

no matter which finite subset  $\widehat{S}_n \subset S_n$  we choose. Suppose now that  $A^{-1}a \cap B^{-1}b \cap S_n \neq \emptyset$  and consider the following cases:

$$(a) \quad A^{-1}a \cup B^{-1}b \subset S_n, \quad (b) \quad A^{-1}a \cup B^{-1}b \not\subset S_n.$$

We will show how to proceed in case (b) (case (a) can be treated in a very similar way). Recall that  $S_n = I^{\mathbb{R}}[(2n-1)\tilde{a}_{n-1}] \times \text{SU}(2)$ , so in view of the definition of  $\mathcal{A}$  one of the following holds:

- (b1)  $A^{-1}a \cup B^{-1}b \subset (l-2, l+1) \times U_i$ ,
- (b2)  $A^{-1}a \cup B^{-1}b \subset (l-1, l+2) \times U_i$ ,
- (b3)  $A^{-1}a \cup B^{-1}b \subset (-l-2, -l+1) \times U_i$ ,
- (b4)  $A^{-1}a \cup B^{-1}b \subset (-l-1, -l+2) \times U_i$ ,

where  $1 \leq i \leq M$  and  $l = -(2n-1)\tilde{a}_{n-1}$ . We will handle case (b1) (the

---

<sup>(24)</sup> The existence of such a sequence can be shown e.g. in the following way. Choose a finite open cover of  $\mathbb{T} \times \text{SU}(2)$  consisting of *continuity sets* (i.e. sets whose boundaries are of zero measure) which are simply connected. It yields a partition of  $\mathbb{T} \times \text{SU}(2)$  into a finite number of open sets, up to a set of measure zero. Using Theorem A.5, for each set of this partition, we can find a homeomorphism to  $(0, 1)^4$  carrying the conditional Haar measure to the Lebesgue measure. Any sequence which is uniformly distributed in  $[0, 1]^4$  yields sequences uniformly distributed in the closure of the elements of the partition. The desired sequence can now be constructed by taking elements from these sequences with frequencies approximating the measures of the sets from the cover.

other cases can be treated in the same way). We have

$$\begin{aligned}
\text{(A.8)} \quad \lambda_{S_n}(A^{-1}a \cap B^{-1}b) &= \frac{\lambda((l-1, l+2) \times \text{SU}(2))}{\lambda(S_n)} \\
&\cdot \lambda_{(l-1, l+2) \times \text{SU}(2)}(A^{-1}a \cap B^{-1}b \cap ((l, l+2) \times \text{SU}(2))) \\
&= \frac{3}{2|l|} \cdot \lambda_{(l-1, l+2) \times \text{SU}(2)}(A^{-1}a \cap B^{-1}b \cap ((l, l+2) \times \text{SU}(2))) \\
&= \frac{3}{2|l|} \cdot \lambda(\psi \circ \psi_{l,i} \circ g_a \circ g \circ \psi_{l_A, i_A}^{-1} \circ \psi^{-1}(C_A) \\
&\quad \cap \psi \circ \psi_{l,i} \circ g_b \circ g \circ \psi_{l_B, i_B}^{-1} \circ \psi^{-1}(C_B) \cap ((1/3, 1) \times (0, 1)^3)) \\
&= \frac{3}{2|l|} \cdot \lambda(\Psi_A(C_A) \cap \Psi_B(C_B) \cap C),
\end{aligned}$$

where

$$\begin{aligned}
\Psi_A &= \psi \circ \psi_{l,i} \circ g_a \circ g \circ \psi_{l_A, i_A}^{-1} \circ \psi^{-1}, \\
\Psi_B &= \psi \circ \psi_{l,i} \circ g_b \circ g \circ \psi_{l_B, i_B}^{-1} \circ \psi^{-1}, \\
C &= (1/3, 1) \times (0, 1)^3.
\end{aligned}$$

Define  $(y_n)_{n \in \mathbb{N}} \subset [0, 1]^4$  by

$$y_n = \begin{cases} \psi \circ \psi_{l,i}(x_{(n+2)/3, l-1}) & \text{if } n \equiv 0 \pmod{3}, \\ \psi \circ \psi_{l,i}(x_{(n+1)/3, l}) & \text{if } n \equiv 1 \pmod{3}, \\ \psi \circ \psi_{l,i}(x_{n/3, l+1}) & \text{if } n \equiv 2 \pmod{3}. \end{cases}$$

Notice that  $(y_n)_{n \in \mathbb{N}}$  is uniformly distributed in  $[0, 1]^4$  and its discrepancy does not depend on  $l \in \mathbb{Z}$ . We have

$$\begin{aligned}
\text{(A.9)} \quad \frac{1}{N} \sum_{n \leq N} \mathbb{1}_{A^{-1}a \cap B^{-1}b \cap ((l, l+2) \times \text{SU}(2))}(x_{n, l-1}) \\
+ \frac{1}{N} \sum_{n \leq N} \mathbb{1}_{A^{-1}a \cap B^{-1}b \cap ((l, l+2) \times \text{SU}(2))}(x_{n, l}) \\
+ \frac{1}{N} \sum_{n \leq N} \mathbb{1}_{A^{-1}a \cap B^{-1}b \cap ((l, l+2) \times \text{SU}(2))}(x_{n, l+1}) \\
= 3 \cdot \frac{1}{3N} \sum_{n=1}^{3N} \mathbb{1}_{\Psi_A(C_A) \cap \Psi_B(C_B) \cap C}(y_n) \rightarrow 3\lambda(\Psi_A(C_A) \cap \Psi_B(C_B) \cap C).
\end{aligned}$$

In view of Lemma A.10 the above convergence is uniform with respect to  $A, B$  and  $a, b$ . It follows from (A.8) and (A.9) that  $\widehat{S}_n$  defined by

$$\widehat{S}_n = S_n \cap \bigcup_{l \in \mathbb{Z}} \bigcup_{n=1}^{n_0} x_{n, l}$$

satisfies (i) provided that  $n_0$  is large enough. ■



*Proof of Proposition A.6(ii).* We claim that the family

$$(A.10) \quad \{f: G \times G \rightarrow \mathbb{R} : f(v, w) = \lambda_{F_n}(Aav \cap Bbw)\}$$

is uniformly equicontinuous. Fix  $\varepsilon > 0$ . Let  $\delta \in (0, \delta_0)$  be so small that for  $C \in \mathcal{A}$  and  $c \in G$  such that  $\lambda(Cc) = \lambda(C) > \varepsilon/2$  we can find a ball of radius  $\delta$  contained in  $Cc$ . Such a choice is possible in view of Lemma A.10. Take  $A, B \in \mathcal{A}$ ,  $a, b \in G$ ,  $n \in \mathbb{N}$  and let  $f(v, w) = \lambda_{F_n}(Aav \cap Bbw)$ . Take  $v_1, v_2, w \in G$  with  $d(v_1, v_2) < \delta$ . We have

$$(A.11) \quad |f(v_1, w) - f(v_2, w)| \leq \lambda((Aav_1 \triangle Aav_2) \cap Bbw).$$

We will consider the following cases:

- (i)  $(Aav_1 \cup Aav_2) \cap Bbw = \emptyset$ ,
- (ii)  $(Aav_1 \cup Aav_2) \cap Bbw \neq \emptyset$ ,  $\lambda(A) \leq \varepsilon/2$ ,
- (iib)  $(Aav_1 \cup Aav_2) \cap Bbw \neq \emptyset$ ,  $\lambda(A) > \varepsilon/2$ .

Using (A.11), in case (i) we obtain  $|f(v_1, w) - f(v_2, w)| = 0$ , and in case (ii),  $|f(v_1, w) - f(v_2, w)| \leq \varepsilon$ . In case (iib), by the choice of  $\delta$ , it is clear that  $Aav_1 \cap Aav_2 \neq \emptyset$ . Therefore (see the definition of  $\tilde{\mathcal{A}}$ ) there exist  $l \in \mathbb{Z}$  and  $1 \leq i \leq M$  such that

$$(Aav_1 \cup Aav_2 \cup Bbw)_{\delta_0} \cup (Aav_1 \cup Aav_2 \cup Bbw) \subset (l-1, l+2) \times U_i.$$

Now we can use the homeomorphism  $\psi_{l,i}^{-1} \circ \psi$  to “transport” the above set to  $[0, 1]^4$ . Using Lemma A.10 and adjusting  $\delta$  if necessary we arrive at (A.10).

Let  $K = (2n-1)\tilde{a}_{n-1}$  and for  $k \in [-K, K-1]$  let  $y_{m,k} = x_m + (k, I)$ . Moreover, for  $1 \leq i \leq M$  let  $(y_{m,k}^i)_{m \in \mathbb{N}}$  be a subsequence of elements of  $(y_{m,k})_{m \in \mathbb{N}}$  which are in  $[k, k+1] \times U_i$  and let  $N_i = \#\{x_{m,k} \in U_i : 1 \leq m \leq N\}$  (notice that this quantity is independent of  $k$ ). Then for a fixed  $w$  we have

$$(A.12) \quad \begin{aligned} & \frac{1}{2KN} \sum_{k=-K}^{K-1} \sum_{m=1}^N f(y_{m,k}, w) \\ &= \frac{1}{2KN} \sum_k \sum_m \sum_l \sum_i \eta_l \times \rho_i(y_{m,k}) f(y_{m,k}, w) \\ &= \frac{1}{2KN} \sum_i \sum_l \sum_{k=l-1}^{l+1} \sum_m \eta_l \times \rho_i(y_{m,k}) f(y_{m,k}, w) \\ &= \frac{1}{2KN} \sum_i \frac{N_i}{N} \sum_l \frac{1}{N_i} \sum_{k=l-1}^{l+1} \sum_{m=1}^{N_i} \underbrace{\eta_l \otimes \rho_i(y_{m,k}^i) f(y_{m,k}, w)}_{f_{l,i,m,k}}. \end{aligned}$$

We deduce that

$$\frac{1}{3N_i} \sum_{k=l-1}^{l+1} \sum_{m=1}^{N_i} f_{l,i,m,k}(y_{m,k}^i) \rightarrow \int f_{l,i,m,k} d\lambda_{(l-1, l+2) \times \text{SU}(2)}$$

and

$$\frac{N_i}{N} \rightarrow \lambda_{\text{SU}(2)}(U_i),$$

and the above convergences are uniform in view of Theorem A.3 and Lemma A.10. Hence the expression in (A.12) converges to

$$\begin{aligned} & \frac{1}{2K} \sum_{i=1}^M \sum_l \lambda_{\text{SU}(2)}(U_i) \cdot 3 \cdot \int f_{l,i,m,k} d\lambda_{(l-1,l+2) \times U_i} \\ &= \frac{1}{2K} \int \sum_l \sum_{i=1}^M 3\lambda_{\text{SU}(2)}(U_i) \eta_l \otimes \rho_i \cdot f(\cdot, w) d\lambda_{(l-1,l+2) \times U_i} = \int f(\cdot, w) d\lambda_{S_n}. \end{aligned}$$

To end the proof it suffices to use Fubini's theorem and take  $\widehat{S}_n := S_n \cap \bigcup_{k \in \mathbb{Z}} \bigcup_{n=1}^{n_0} x_{n,k}$  for  $n_0$  large enough. Notice that this is consistent with the final choice of  $\widehat{S}_n$  in the proof of Proposition A.6(i). ■

REMARK A.12. Without loss of generality, in Proposition A.6 we may assume that  $F_n \in \mathcal{A}$  for  $n \in \mathbb{N}$ .

REMARK A.13. The assertion of Proposition A.6 remains true if instead of  $A, B \in \mathcal{A}$  we consider sets  $A', B'$  which are finite unions of translations of elements from  $\mathcal{A}$ , provided that the number of elements in the union is bounded for each  $n$ .

**Acknowledgments.** Research supported by Narodowe Centrum Nauki grant DEC-2011/03/B/ST1/00407.

## References

- [1] L. M. Abramov and V. A. Rokhlin, *Entropy of a skew product of mappings with invariant measure*, Vestnik Leningrad. Univ. 17 (1962), 5–13 (in Russian).
- [2] W. Ambrose, *Representation of ergodic flows*, Ann. of Math. (2) 42 (1941), 723–739.
- [3] W. Ambrose and S. Kakutani, *Structure and continuity of measurable flows*, Duke Math. J. 9 (1942), 25–42.
- [4] T. Austin, *Pleasant extensions retaining algebraic structure. I*, arXiv.org:0905.0518 (2009).
- [5] J. R. Blum and N. Friedman, *On commuting transformations and roots*, Proc. Amer. Math. Soc. 17 (1966), 1370–1374.
- [6] W. Bułatek, M. Lemańczyk, and E. Lesigne, *On the filtering problem for stationary random  $\mathbb{Z}^2$ -fields*, IEEE Trans. Inform. Theory 51 (2005), 3586–3593.
- [7] R. V. Chacon, *Transformations having continuous spectrum*, J. Math. Mech. 16 (1966), 399–415.
- [8] R. V. Chacon, *A geometric construction of measure preserving transformations*, in: Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, CA, 1965/66), Vol. II: Contributions to Probability Theory, Part 2, Univ. of California Press, CA, 1967, 335–360.

- [9] R. V. Chacon, *Weakly mixing transformations which are not strongly mixing*, Proc. Amer. Math. Soc. 22 (1969), 559–562.
- [10] I. P. Cornfeld, S. V. Fomin, and Y. G. Sinai, *Ergodic Theory*, Grundlehren Math. Wiss. 245, Springer, New York, 1982.
- [11] A. I. Danilenko, *On simplicity concepts for ergodic actions*, J. Anal. Math. 102 (2007), 77–117.
- [12] A. I. Danilenko, *(C, F)-actions in ergodic theory*, in: Geometry and Dynamics of Groups and Spaces, Progr. Math. 265, Birkhäuser, Basel, 2008, 325–351.
- [13] A. I. Danilenko, *Mixing actions of the Heisenberg group*, Ergodic Theory Dynam. Systems 34 (2014), 1142–1167.
- [14] T. de la Rue, *Notes on Austin’s multiple ergodic theorem*, arXiv:0907.0538 (2009).
- [15] T. de la Rue et J. de Sam Lazaro, *Une transformation générique peut être insérée dans un flot*, Ann. Inst. H. Poincaré Probab. Statist. 39 (2003), 121–134.
- [16] A. del Junco, *A simple measure-preserving transformation with trivial centralizer*, Pacific J. Math. 79 (1978), 357–362.
- [17] A. del Junco, *A simple map with no prime factors*, Israel J. Math. 104 (1998), 301–320.
- [18] A. del Junco and M. Lemańczyk, *Simple systems are disjoint from Gaussian systems*, Studia Math. 133 (1999), 249–256.
- [19] A. del Junco and M. Lemańczyk, *Joinings of distally simple automorphisms*, unpublished preprint, 2005.
- [20] A. del Junco and D. Rudolph, *On ergodic actions whose self-joinings are graphs*, Ergodic Theory Dynam. Systems 7 (1987), 531–557.
- [21] A. del Junco and D. J. Rudolph, *A rank-one, rigid, simple, prime map*, Ergodic Theory Dynam. Systems 7 (1987), 229–247.
- [22] B. Fayad and A. Kanigowski, *On multiple mixing for a class of conservative surface flows*, preprint.
- [23] S. Ferenczi, *Systèmes de rang un gauche*, Ann. Inst. H. Poincaré Probab. Statist. 21 (1985), 177–186.
- [24] K. Frączek and M. Lemańczyk, *On mild mixing of special flows over irrational rotations under piecewise smooth functions*, Ergodic Theory Dynam. Systems 26 (2006), 719–738.
- [25] K. Frączek and M. Lemańczyk, *Ratner’s property and mild mixing for special flows over two-dimensional rotations*, J. Modern Dynam. 4 (2010), 609–635.
- [26] H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Math. Systems Theory 1 (1967), 1–49.
- [27] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton Univ. Press, Princeton, NJ, 1981.
- [28] E. Glasner, B. Host, and D. Rudolph, *Simple systems and their higher order self-joinings*, Israel J. Math. 78 (1992), 131–142.
- [29] S. Glasner, *Quasifactors in ergodic theory*, Israel J. Math. 45 (1983), 198–208.
- [30] F. Hahn and W. Parry, *Some characteristic properties of dynamical systems with quasi-discrete spectra*, Math. Systems Theory 2 (1968), 179–190.
- [31] P. R. Halmos, *Square roots of measure preserving transformations*, Amer. J. Math. 64 (1942), 153–166.
- [32] P. R. Halmos, *Lectures on Ergodic Theory*, Publ. Math. Soc. Japan 3, Math. Soc. Japan, Tokyo, 1956.
- [33] P. R. Halmos, *Lectures on Ergodic Theory*, Chelsea, New York, 1960.

- [34] E. Hlawka, *Discrepancy and Riemann integration*, in: Studies in Pure Mathematics (Presented to Richard Rado), Academic Press, London, 1971, 121–129.
- [35] A. Kanigowski, *Ratner's property for special flows over irrational rotations under functions of bounded variation*, Ergodic Theory Dynam. Systems (2013) (online).
- [36] A. Katok, *Combinatorial Constructions in Ergodic Theory and Dynamics*, Univ. Lecture Ser. 30, Amer. Math. Soc., Providence, RI, 2003.
- [37] J. L. F. King, *The generic transformation has roots of all orders*, Colloq. Math. 84/85 (2000), 521–547.
- [38] A. V. Kochergin, *The homology of functions over dynamical systems*, Dokl. Akad. Nauk SSSR 231 (1976), 795–798 (in Russian).
- [39] U. Krengel und H. Michel, *Über Wurzeln ergodischer Transformationen*, Math. Z. 96 (1967), 50–57.
- [40] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Pure Appl. Math., Wiley-Interscience, New York, 1974.
- [41] J. Kułaga-Przymus and F. Parreau, *Disjointness properties for Cartesian products of weakly mixing systems*, Colloq. Math. 128 (2012), 153–177.
- [42] J. Lamperti, *On flows of measure-preserving transformations*, Michigan Math. J. 4 (1957), 161–164.
- [43] M. Lemańczyk and M. K. Mentzen, *Compact subgroups in the centralizer of natural factors of an ergodic group extension of a rotation determine all factors*, Ergodic Theory Dynam. Systems 10 (1990), 763–776.
- [44] M. Lemańczyk and F. Parreau, *Rokhlin extensions and lifting disjointness*, Ergodic Theory Dynam. Systems 23 (2003), 1525–1550.
- [45] M. Lemańczyk, F. Parreau, and E. Roy, *Joining primeness and disjointness from infinitely divisible systems*, Proc. Amer. Math. Soc. 139 (2011), 185–199.
- [46] M. Lemańczyk, F. Parreau, and J.-P. Thouvenot, *Gaussian automorphisms whose ergodic self-joinings are Gaussian*, Fund. Math. 164 (2000), 253–293.
- [47] M. Lemańczyk, J.-P. Thouvenot, and B. Weiss, *Relative discrete spectrum and joinings*, Monatsh. Math. 137 (2002), 57–75.
- [48] E. Lindenstrauss, *Pointwise theorems for amenable groups*, Invent. Math. 146 (2001), 259–295.
- [49] M. K. Mentzen, *Ergodic properties of group extensions of dynamical systems with discrete spectra*, Studia Math. 101 (1991), 19–31.
- [50] H. Michel, *Wurzeln vollergodischer Automorphismen mit quasidiskretem Spektrum. I*, Z. Wahrsch. Verw. Gebiete 9 (1968), 323–340.
- [51] H. Michel, *Vollergodische Automorphismen mit quasidiskretem Spektrum und Wurzeln jeder rationalen Ordnung*, Wiss. Beiträge Martin-Luther-Univ. Halle-Wittenberg M1 (1968/1969), 123–128.
- [52] H. Michel, *Wurzeln vollergodischer Automorphismen mit quasidiskretem Spektrum. II*, Z. Wahrsch. Verw. Gebiete 16 (1970), 321–335.
- [53] J. Neveu, *Processus aléatoires gaussiens*, Sémin. Math. Supér. 34, Presses Univ. de Montréal, Montréal, 1968.
- [54] D. Newton, *On canonical factors of ergodic dynamical systems*, J. London Math. Soc. (2) 19 (1979), 129–136.
- [55] D. Ornstein, *Bernoulli shifts with the same entropy are isomorphic*, Adv. Math. 4 (1970), 337–352.
- [56] D. Ornstein, *Two Bernoulli shifts with infinite entropy are isomorphic*, Adv. Math. 5 (1970), 339–348.

- [57] D. S. Ornstein, *Imbedding Bernoulli shifts in flows*, in: Contributions to Ergodic Theory and Probability (Columbus, OH, 1970), Springer, Berlin, 1970, 178–218.
- [58] D. S. Ornstein, *On the root problem in ergodic theory*, in: Proc. Sixth Berkeley Sympos. of Mathematical Statistics and Probability (Berkeley, CA, 1970/1971), Vol. II: Probability Theory, CA, Univ. of California Press, 1972, 347–356.
- [59] D. S. Ornstein, *The isomorphism theorem for Bernoulli flows*, Adv. Math. 10 (1973), 124–142.
- [60] D. S. Ornstein, *A  $K$  automorphism with no square root and Pinsker's conjecture*, Adv. Math. 10 (1973), 89–102.
- [61] M. Osikawa, *Centralizer of an ergodic measure preserving transformation*, Publ. RIMS Kyoto Univ. 18 (1982), 35–48.
- [62] J. C. Oxtoby and S. M. Ulam, *Measure-preserving homeomorphisms and metrical transitivity*, Ann. of Math. (2) 42 (1941), 874–920.
- [63] W. Parry, *Zero entropy of distal and related transformations*, in: Topological Dynamics (Ft. Collins, CO, 1967), Benjamin, New York, 1968, 383–389.
- [64] M. Ratner, *Factors of horocycle flows*, Ergodic Theory Dynam. Systems 2 (1982), 465–489.
- [65] M. Ratner, *Rigidity of horocycle flows*, Ann. of Math. (2) 115 (1982), 597–614.
- [66] M. Ratner, *Horocycle flows, joinings and rigidity of products*, Ann. of Math. (2) 118 (1983), 277–313.
- [67] D. Rudolph, *Two nonisomorphic  $K$ -automorphisms all of whose powers beyond one are isomorphic*, Israel J. Math. 27 (1977), 277–298.
- [68] D. J. Rudolph, *Two nonisomorphic  $K$ -automorphisms with isomorphic squares*, Israel J. Math. 23 (1976), 274–287.
- [69] D. J. Rudolph, *An example of a measure preserving map with minimal self-joinings, and applications*, J. Anal. Math. 35 (1979), 97–122.
- [70] V. V. Ryzhikov and J.-P. Thouvenot, *Disjointness, divisibility, and quasi-simplicity of measure-preserving actions*, Funktsional. Anal. i Prilozhen. 40 (2006), no. 3, 85–89 (in Russian).
- [71] V. V. Ryzhikov and S. V. Tikhonov, *Generic  $\mathbb{Z}^n$ -action can be embedded only in injective  $\mathbb{R}^n$ -actions*, Mat. Zametki 79 (2006), 925–930 (in Russian).
- [72] A. Siemaszko, *On group extensions of 2-fold simple ergodic actions*, Studia Math. 110 (1994), 53–64.
- [73] A. Sokhet, *Les actions approximativement transitives dans la théorie ergodique*, thèse de doctorat, Univ. Paris VII, 2007.
- [74] A. M. Stepin and A. M. Eremenko, *Nonuniqueness of an inclusion in a flow and the vastness of a centralizer for a generic measure-preserving transformation*, Mat. Sb. 195 (2004), no. 12, 95–108 (in Russian).
- [75] J.-P. Thouvenot, *Some properties and applications of joinings in ergodic theory*, in: Ergodic Theory and Its Connections with Harmonic Analysis (Alexandria, 1993), London Math. Soc. Lecture Note Ser. 205, Cambridge Univ. Press, Cambridge, 1995, 207–235.
- [76] S. V. Tikhonov, *The generic action of the group  $\mathbb{Z}^d$  can be embedded into the action of the group  $\mathbb{R}^d$* , Dokl. Akad. Nauk 391 (2003), 26–28 (in Russian).
- [77] S. V. Tikhonov, *Embeddings of lattice actions in flows with multidimensional time*, Mat. Sb. 197 (2006), no. 1, 97–132 (in Russian).
- [78] W. A. Veech, *A criterion for a process to be prime*, Monatsh. Math. 94 (1982), 335–341.
- [79] D. Witte, *Rigidity of some translations on homogeneous spaces*, Invent. Math. 81 (1985), 1–27.

- [80] R. J. Zimmer, *Ergodic actions with generalized discrete spectrum*, Illinois J. Math. 20 (1976), 555–588.
- [81] R. J. Zimmer, *Extensions of ergodic group actions*, Illinois J. Math. 20 (1976), 373–409.
- [82] R. J. Zimmer, *Induced and amenable ergodic actions of Lie groups*, Ann. Sci. École Norm. Sup. (4) 11 (1978), 407–428.

Joanna Kułaga-Przymus  
Institute of Mathematics  
Polish Academy of Sciences  
Śniadeckich 8  
00-656 Warszawa, Poland  
and  
Faculty of Mathematics and Computer Science  
Nicolaus Copernicus University  
Chopina 12/18  
87-100 Toruń, Poland  
E-mail: joanna.kulaga@gmail.com

*Received 12 April 2013;  
in revised form 2 September 2014*