STRICTLY ERGODIC MODELS UNDER FACE AND PARALLELEPIPED GROUP ACTIONS

WEN HUANG, SONG SHAO, AND XIANGDONG YE

Abstract. The Jewett-Krieger Theorem states that each ergodic system has a strictly ergodic topological model. In this article, we show that for an ergodic system one may require more properties on its strictly ergodic model. For example, the orbit closure of points in diagonal under face transforms may be also strictly ergodic. As an application, we show the pointwise convergence of ergodic averages along cubes, which was firstly proved by Assani [1].

1. Introduction

In the introduction we will state the main results of the paper and give main ideas of proofs.

1.1. **Main results.** Throughout this paper, by a *topological dynamical system* (t.d.s. for short) we mean a pair (X,T), where X is a compact metric space and T is a homeomorphism from X to itself. A *measurable system* (m.p.t. for short) is a quadruple $(X,\mathcal{X},\mu,T)$, where $(X,\mathcal{X},\mu)$ is a Lebesgue probability space and $T:X\to X$ is an invertible measure preserving transformation.

Let $(X,\mathcal{X},\mu,T)$ be an ergodic m.p.t. We say that $(\hat{X},\hat{T})$ is a *topological model* (or just a *model*) for $(X,\mathcal{X},\mu,T)$ if $(\hat{X},\hat{T})$ is a t.d.s. and there exists an invariant probability measure $\hat{\mu}$ on the Borel σ-algebra $\mathcal{B}(\hat{X})$ such that the systems $(X,\mathcal{X},\mu,T)$ and $(\hat{X},\mathcal{B}(\hat{X}),\hat{\mu},\hat{T})$ are measure theoretically isomorphic.

The well-known Jewett-Krieger's theorem [14, 15] states that every ergodic system has a strictly ergodic model. We note that one can add some additional properties to the topological model. For example, in [16] Lehrer showed that the strictly ergodic model can be required to be a topological (strongly) mixing system in addition.

Let $(\hat{X},\hat{T})$ be a t.d.s. Write $(x,\dots,x)$ (2^d times) as $x^{[d]}$. Let $\mathcal{F}^{[d]},\mathcal{G}^{[d]}$ and $\mathbf{Q}^{[d]}(\hat{X})$ be the face group of dimension d, the parallelepiped group of dimension d and the dynamical parallelepiped of dimension d respectively (see Section 2 for definitions). The orbit closure of $x^{[d]}$ under the face group action will be denote by $\overline{\mathcal{F}^{[d]}}(x^{[d]})$. It was shown by Shao and Ye [18] that if $(\hat{X},\hat{T})$ is minimal then $(\overline{\mathcal{F}^{[d]}}(x^{[d]}),\mathcal{F}^{[d]})$ is minimal for all $x\in\hat{X}$ and $(\mathbf{Q}^{[d]}(\hat{X}),\mathcal{G}^{[d]})$ is minimal.

Date: July 03, 2014.

2000 *Mathematics Subject Classification.* Primary: 37B05, 37A05.

Key words and phrases. ergodic averages, model, cubes, face transformations.

Huang is partially supported by NNSF for Distinguished Young School (11225105), Shao and Ye is partially supported by NNSF of China (11171320), and Huang and Ye is partially supported by NNSF of China (11371339).

In this paper we will strengthen Jewett-Krieger's theorem in another direction. Namely, we have the following theorem.

Theorem A: *Let $(X,\mathcal{X},\mu,T)$ be an ergodic m.p.t. and $d\in\mathbb{N}$. Then*

1. *it has a strictly ergodic model $(\hat{X},\hat{T})$ such that $(\overline{\mathcal{F}^{[d]}}(x^{[d]}),\mathcal{F}^{[d]})$ is strictly ergodic for all $x\in\hat{X}$.*
2. *it has a strictly ergodic model $(\hat{X},\hat{T})$ such that $(\mathbf{Q}^{[d]}(\hat{X}),\mathcal{G}^{[d]})$ is strictly ergodic.*

Note that we have formulas to compute the unique measure in Theorems A. Particularly, when $(X,\mathcal{X},\mu,T)$ is weakly mixing, the unique measure is nothing but the product measure.

As an application, we have

Theorem B: *Let $(X,\mathcal{X},\mu,T)$ be an ergodic m.p.t. and $d\in\mathbb{N}$. Then*

(1) *for functions $f_\epsilon\in L^\infty(\mu)$, $\epsilon\in\{0,1\}^d,\epsilon\neq(0,\dots,0)$, the averages*

$$\frac{1}{N^d}\sum_{\mathbf{n}\in\{0,1,\dots,N-1\}^d}\prod_{(0,\dots,0)\neq\epsilon\in\{0,1\}^d}f_\epsilon(T^{\mathbf{n}\cdot\epsilon}x) \tag{1.1}$$

converge μ a.e..

(2) *for functions $f_\epsilon\in L^\infty(\mu)$, $\epsilon\in\{0,1\}^d$, the averages*

$$\frac{1}{N^{d+1}}\sum_{\substack{\mathbf{n}\in\{0,1,\dots,N-1\}^d\\ n\in\{0,1,\dots,N-1\}}}\prod_{\epsilon\subset[d]}f_\epsilon(T^{n+\mathbf{n}\cdot\epsilon}x) \tag{1.2}$$

converge μ a.e..

Remark 1.1. The study of the limiting behavior of the averages along cubes was initiated by Bergelson in [2], where convergence in $L^2(\mu)$ was shown in dimension 2. Bergelson's result was later extended by Host and Kra for cubic averages of an arbitrary dimension d in [9]. More recently in [1], Assani established pointwise convergence for cubic averages of an arbitrary dimension d. Chu and Franzikinakis [4] extended the result to a very general case, i.e. they showed that for measure preserving transformations $T_\epsilon: X\to X$, functions $f_\epsilon\in L^\infty(\mu)$, $(0,\dots,0)\neq\epsilon\in\{0,1\}^d$, the averages

$$\frac{1}{N^d}\sum_{\mathbf{n}\in[0,N-1]^d}\prod_{(0,\dots,0)\neq\epsilon\in\{0,1\}^d}f_\epsilon(T_\epsilon^{\mathbf{n}\cdot\epsilon}x)$$

converge μ a.e..

Using the similar methods as in this paper, we prove in [13] that for an ergodic system $(X,\mathcal{X},\mu,T)$, $d\in\mathbb{N}$, $f_1,\dots,f_d\in L^\infty(\mu)$, the averages

$$\frac{1}{N^2}\sum_{(n,m)\in[0,N-1]^2}f_1(T^nx)f_2(T^{n+m}x)\dots f_d(T^{n+(d-1)m}x)$$

converge μ a.e.

In the same paper[13], for distal systems we answer positively the question if the multiple ergodic averages converge a.e. That is, we show that if $(X, \mathcal{X}, \mu, T)$ is an ergodic distal system, and $f_1, \ldots, f_d \in L^\infty(\mu)$, then multiple ergodic averages

$$\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^n x) \ldots f_d(T^{dn} x)$$

converge μ a.e..

1.2. **Main ideas of the proofs.** Now we describe the main ideas and ingredients in the proof of Theorem A. The first fact we face is that for an ergodic m.p.t. $(X, \mathcal{X}, \mu, T)$, not every strictly ergodic model is its $\mathcal{F}^{[d]}$-strictly ergodic model. For example, let $(X, \mathcal{X}, \mu, T)$ be a Kronecker system. By Jewett-Krieger' Theorem, we may assume that (X, T) is a topologically weakly mixing minimal system and strictly ergodic. By [18, Theorem 3.11] $(\overline{\mathcal{F}^{[d]}}(x^{[d]}), \mathcal{F}^{[d]})$ is minimal for all $x \in X$ and $\overline{\mathcal{F}^{[d]}}(x^{[d]}) = \{x\} \times X_*^{[d]}$. It is easy to see that $\delta_x \times \mu^{\bigotimes 2^d - 1}$ and $\mu_*^{[d]}$ are two different invariant measures on it (see Section 2 for the definitions). This indicates that to obtain Theorem A, Jewett-Krieger' Theorem is not enough for our purpose. Fortunately, we find that Weiss's Theorem [19] is a right tool.

Precisely, let $\pi_d : X \to Z_d$ be the factor map from X to its d-step nilfactor Z_d. By definition, Z_d may be regarded as a topological system in the natural way. By Weiss's Theorem there is a uniquely ergodic model $(\hat{X}, \hat{\mathcal{X}}, \hat{\mu}, \hat{T})$ for $(X, \mathcal{X}, \mu, T)$ and a factor map $\hat{\pi_d} : \hat{X} \to Z_d$ which is a model for $\pi_d : X \to Z_d$.

$$\begin{array}{ccc} X & \xrightarrow{\phi} & \hat{X} \\ {\scriptstyle \pi_d}\downarrow & & \downarrow{\scriptstyle \hat{\pi}_d} \\ Z_d & \longrightarrow & Z_d \end{array}$$

We then show (though it is difficult) that $(\hat{X}, \hat{T})$ is what we need. To do this we heavily use the theory of joinings (for a reference, see [7]) and some facts related to d-step nilsystems.

Once Theorem A is proven, Theorem B will follow by an argument using some well known theorems related to pointwise convergence for $\mathbb{Z}^d$ actions by and for uniquely ergodic systems.

1.3. **Organization of the paper.** In Section 2, we give basic notions and facts about dynamical parallelepipeds and characteristic factors. In Section 3 we define $\mathcal{F}$ and $\mathcal{G}$-strictly ergodic models and prove that each ergodic system has $\mathcal{F}$ and $\mathcal{G}$-strictly ergodic model. Moreover, we build the connection between $\mathcal{F}$ and $\mathcal{G}$-strictly ergodic models with pointwise convergence of averages along cubes and faces, and deduce the existence of the limit of the averages. In the last section, we prove Theorem A.

2. Dynamical parallelepipeds and characteristic factors

In this section we introduce basic knowledge about dynamical parallelepipeds and characteristic factors. For more details, see [9, 10, 11] etc.

2.1. **Ergodic theory and topological dynamics.** In this subsection we introduce some basic notions in ergodic theory and topological dynamics. For more information, see Appendix of [13].

2.1.1. *Measurable systems.* For a m.p.t. $(X,\mathcal{X},\mu,T)$ we write $\mathcal{I}=\mathcal{I}(T)$ for the σ-algebra $\{A\in\mathcal{X}:T^{-1}A=A\}$ of invariant sets. A m.p.t. is *ergodic* if all the T-invariant sets have measure either 0 or 1. $(X,\mathcal{X},\mu,T)$ is *weakly mixing* if the product system $(X\times X,\mathcal{X}\times\mathcal{X},\mu\times\mu,T\times T)$ is erdogic.

A *homomorphism* from m.p.t. $(X,\mathcal{X},\mu,T)$ to $(Y,\mathcal{Y},\nu,S)$ is a measurable map $\pi:X_0\to Y_0$, where X_0 is a T-invariant subset of X and Y_0 is an S-invariant subset of Y, both of full measure, such that $\pi_*\mu=\mu\circ\pi^{-1}=\nu$ and $S\circ\pi(x)=\pi\circ T(x)$ for $x\in X_0$. When we have such a homomorphism we say that $(Y,\mathcal{Y},\nu,S)$ is a *factor* of $(X,\mathcal{X},\mu,T)$. If the factor map $\pi:X_0\to Y_0$ can be chosen to be bijective, then we say that $(X,\mathcal{X},\mu,T)$ and $(Y,\mathcal{Y},\nu,S)$ are *(measure theoretically) isomorphic*. A factor can be characterized (modulo isomorphism) by $\pi^{-1}(\mathcal{Y})$, which is a T-invariant sub- σ-algebra of $\mathcal{X}$, and conversely any T-invariant sub-σ-algebra of $\mathcal{X}$ defines a factor. By a classical result abuse of terminology we denote by the same letter the σ-algebra $\mathcal{Y}$ and its inverse image by π.

2.1.2. *Topological dynamical systems.* A t.d.s. (X,T) is *transitive* if there exists some point $x\in X$ whose orbit $\mathcal{O}(x,T)=\{T^nx:n\in\mathbb{Z}\}$ is dense in X and we call such a point a *transitive point*. The system is *minimal* if the orbit of any point is dense in X. (X,T) is *topologically weakly mixing* if the product system $(X\times X,T\times T)$ is transitive.

A *factor* of a t.d.s. (X,T) is another t.d.s. (Y,S) such that there exists a continuous and onto map $\phi:X\to Y$ satisfying $S\circ\phi=\phi\circ T$. In this case, (X,T) is called an *extension* of (Y,S). The map ϕ is called a *factor map*.

2.1.3. We also make use of a more general definition of a measurable or topological system. That is, instead of just a single transformation T, we consider commuting transformations $T_1,\dots,T_k$ of X or a countable abelian group of transformations. We summarize some basic definitions and properties of systems in the classical setting of one transformation. Extensions to the general case are straightforward.

2.1.4. $M(X)$ *and* $M_T(X)$. For a t.d.s. (X,T), denote by $M(X)$ the set of all probability measure on X. Let $M_T(X)=\{\mu\in M(X):T_*\mu=\mu\circ T^{-1}=\mu\}$ be the set of all T-invariant measure of X. It is well known that $M_T(X)\neq\emptyset$.

Definition 2.1. A t.d.s. (X,T) is called *uniquely ergodic* if there is a unique T-invariant probability measure on X. It is called *strictly ergodic* if it is uniquely ergodic and minimal.

2.1.5. *Uniquely ergodic systems.* In this subsection we give some conditions for unique ergodicity under $\mathbb{Z}^d$ actions ($d\in\mathbb{N}$).

Theorem 2.2. *Let (X,Γ) be a topological system, where $\Gamma=\mathbb{Z}^d$. The following conditions are equivalent.*

(1) *(X,Γ) is uniquely ergodic.*

(2) *For every continuous function $f \in C(X)$ the sequence of functions*

$$\mathbb{A}_N f(x) = \frac{1}{N^d} \sum_{\gamma \in [0,N-1]^d} f(\gamma x) \tag{2.1}$$

converges uniformly to a constant function.

(3) *For every continuous function $f \in C(X)$ the sequence of functions $\mathbb{A}_N f(x)$ converges pointwise to a constant function.*

(4) *There exists a $\mu \in M_\Gamma(X)$ such that for all continuous function $f \in C(X)$ and all $x \in X$ the sequence of functions*

$$\mathbb{A}_N f(x) \longrightarrow \int f \, d\mu, \ N \to \infty. \tag{2.2}$$

2.2. **Topological models.**

Definition 2.3. Let $(X, \mathcal{X}, \mu, T)$ be an ergodic dynamical system. We say that the system $(\hat{X}, \hat{\mathcal{X}}, \hat{\mu}, \hat{T})$ is a *topological model* (or just a *model*) for $(X, \mathcal{X}, \mu, T)$ if $(\hat{X}, \hat{T})$ is a topological system, $\hat{\mu} \in M_T(\hat{X})$ and the the systems $(X, \mathcal{X}, \mu, T)$ and $(\hat{X}, \hat{\mathcal{X}}, \hat{\mu}, \hat{T})$ are measure theoretically isomorphic.

It is well known that each system has topological model [5]. Weiss [20] showed the following surprising result: There exists a minimal metric dynamical system (X, T) with the property that for every ergodic probability measure preserving system $(Y, \mathcal{Y}, \nu, S)$ there exists a T-invariant Borel probability measure μ on X such that the systems $(Y, \mathcal{Y}, \nu, S)$ and $(X, \mathcal{B}(X), \mu, T)$ are measure theoretically isomorphic.

Similarly we say that $\hat{\pi} : \hat{X} \to \hat{Y}$ is a *topological model* for $\pi : (X, \mathcal{X}, \mu, T) \to (Y, \mathcal{Y}, \nu, S)$ when $\hat{\pi}$ is a topological factor map and there exist measure theoretical isomorphisms ϕ and ψ such that the diagram

$$\begin{array}{ccc} X & \xrightarrow{\ \phi\ } & \hat{X} \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \hat{\pi}} \\ Y & \xrightarrow{\ \psi\ } & \hat{Y} \end{array}$$

is commutative, i.e. $\hat{\pi}\phi = \psi\pi$.

Here is the famous Jewett-Krieger Theorem:

Theorem 2.4 (Jewett-Krieger)**.** [14, 15] *Every ergodic systems has a uniquely ergodic model.*

B. Weiss generalized this theorem to the relative case.

Theorem 2.5 (B. Weiss)**.** [19] *If $\pi : (X, \mathcal{X}, \mu, T) \to (Y, \mathcal{Y}, \nu, S)$ is a factor map with $(X, \mathcal{X}, \mu, T)$ ergodic and $(\hat{Y}, \hat{\mathcal{Y}}, \hat{\nu}, \hat{S})$ is a uniquely ergodic model for $(Y, \mathcal{Y}, \nu, S)$, then there is a uniquely ergodic model $(\hat{X}, \hat{\mathcal{X}}, \hat{\mu}, \hat{T})$ for $(X, \mathcal{X}, \mu, T)$ and a factor map $\hat{\pi} : \hat{X} \to \hat{Y}$ which is a model for $\pi : X \to Y$.*

2.3. **Cubes and faces.**

2.3.1. Let X be a set, let $d \geq 1$ be an integer, and write $[d] = \{1, 2, \ldots, d\}$. We view $\{0,1\}^d$ in one of two ways, either as a sequence $\epsilon = \epsilon_1 \ldots \epsilon_d$ of $0'$s and $1'$s, or as a subset of $[d]$. A subset ϵ corresponds to the sequence $(\epsilon_1, \ldots, \epsilon_d) \in \{0,1\}^d$ such that $i \in \epsilon$ if and only if $\epsilon_i = 1$ for $i \in [d]$. For example, $\mathbf{0} = (0, 0, \ldots, 0) \in \{0,1\}^d$ is the same to $\emptyset \subset [d]$.

Let $V_d = \{0,1\}^d = [d]$ and $V_d^* = V_d \setminus \{\mathbf{0}\} = V_d \setminus \{\emptyset\}$. If $\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{Z}^d$ and $\epsilon \in \{0,1\}^d$, we define

$$\mathbf{n} \cdot \epsilon = \sum_{i=1}^{d} n_i \epsilon_i.$$

If we consider ϵ as $\epsilon \subset [d]$, then $\mathbf{n} \cdot \epsilon = \sum_{i \in \epsilon} n_i$.

2.3.2. We denote X^{2^d} by $X^{[d]}$. A point $\mathbf{x} \in X^{[d]}$ can be written in one of two equivalent ways, depending on the context:

$$\mathbf{x} = (x_\epsilon : \epsilon \in \{0,1\}^d) = (x_\epsilon : \epsilon \subset [d]).$$

Hence $x_\emptyset = x_{\mathbf{0}}$ is the first coordinate of $\mathbf{x}$. As examples, points in $X^{[2]}$ are like

$$(x_{00}, x_{10}, x_{01}, x_{11}) = (x_\emptyset, x_{\{1\}}, x_{\{2\}}, x_{\{1,2\}}).$$

For $x \in X$, we write $x^{[d]} = (x, x, \ldots, x) \in X^{[d]}$. The diagonal of $X^{[d]}$ is $\Delta^{[d]} = \{x^{[d]} : x \in X\}$. Usually, when $d = 1$, denote diagonal by Δ_X or Δ instead of $\Delta^{[1]}$.

A point $\mathbf{x} \in X^{[d]}$ can be decomposed as $\mathbf{x} = (\mathbf{x}', \mathbf{x}'')$ with $\mathbf{x}', \mathbf{x}'' \in X^{[d-1]}$, where $\mathbf{x}' = (x_{\epsilon 0} : \epsilon \in \{0,1\}^{d-1})$ and $\mathbf{x}'' = (x_{\epsilon 1} : \epsilon \in \{0,1\}^{d-1})$. We can also isolate the first coordinate, writing $X_*^{[d]} = X^{2^d - 1}$ and then writing a point $\mathbf{x} \in X^{[d]}$ as $\mathbf{x} = (x_\emptyset, \mathbf{x}_*)$, where $\mathbf{x}_* = (x_\epsilon : \epsilon \neq \emptyset) \in X_*^{[d]}$.

2.3.3. The *faces* of dimension r of a point in $\mathbf{x} \in X^{[d]}$ are defined as follows. Let $J \subset [d]$ with $|J| = d - r$ and $\xi \in \{0,1\}^{d-r}$. The elements $(x_\epsilon : \epsilon \in \{0,1\}^d, \epsilon_J = \xi)$ of $X^{[r]}$ are called *faces of dimension r* of $\mathbf{x}$, where $\epsilon_J = (\epsilon_i : i \in J)$. Thus any face of dimension r defines a natural projection from $X^{[d]}$ to $X^{[r]}$, and we call this the projection along this face.

2.4. **Dynamical parallelepipeds.**

Definition 2.6. Let (X, T) be a topological dynamical system and let $d \geq 1$ be an integer. We define $\mathbf{Q}^{[d]}(X)$ to be the closure in $X^{[d]}$ of elements of the form

$$(T^{\mathbf{n} \cdot \epsilon} x = T^{n_1 \epsilon_1 + \ldots + n_d \epsilon_d} x : \epsilon = (\epsilon_1, \ldots, \epsilon_d) \in \{0,1\}^d),$$

where $\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{Z}^d$ and $x \in X$. When there is no ambiguity, we write $\mathbf{Q}^{[d]}$ instead of $\mathbf{Q}^{[d]}(X)$. An element of $\mathbf{Q}^{[d]}(X)$ is called a (dynamical) *parallelepiped of dimension d*.

As examples, $\mathbf{Q}^{[2]}$ is the closure in $X^{[2]} = X^4$ of the set

$$\{(x, T^m x, T^n x, T^{n+m} x) : x \in X, m, n \in \mathbb{Z}\}$$

and $\mathbf{Q}^{[3]}$ is the closure in $X^{[3]} = X^8$ of the set

$$\{(x, T^m x, T^n x, T^{m+n} x, T^p x, T^{m+p} x, T^{n+p} x, T^{m+n+p} x) : x \in X, m, n, p \in \mathbb{Z}\}.$$

Definition 2.7. Let $\phi : X \to Y$ and $d \in \mathbb{N}$. Define $\phi^{[d]} : X^{[d]} \to Y^{[d]}$ by $(\phi^{[d]}\mathbf{x})_\epsilon = \phi x_\epsilon$ for every $\mathbf{x} \in X^{[d]}$ and every $\epsilon \subset [d]$. Let (X, T) be a system and $d \geq 1$ be an integer. The *diagonal transformation* of $X^{[d]}$ is the map $T^{[d]}$.

Definition 2.8. *Face transformations* are defined inductively as follows: Let $T^{[0]} = T$, $T_1^{[1]} = \mathrm{id} \times T$. If $\{T_j^{[d-1]}\}_{j=1}^{d-1}$ is defined already, then set

$$
\begin{aligned}
T_j^{[d]} &= T_j^{[d-1]} \times T_j^{[d-1]},\ j \in \{1, 2, \ldots, d-1\}, \\
T_d^{[d]} &= \mathrm{id}^{[d-1]} \times T^{[d-1]}.
\end{aligned}
\tag{2.3}
$$

The *face group* of dimension d is the group $\mathcal{F}^{[d]}(X)$ of transformations of $X^{[d]}$ spanned by the face transformations. The *cube group* or *parallelepiped group* of dimension d is the group $\mathcal{G}^{[d]}(X)$ spanned by the diagonal transformation and the face transformations. We often write $\mathcal{F}^{[d]}$ and $\mathcal{G}^{[d]}$ instead of $\mathcal{F}^{[d]}(X)$ and $\mathcal{G}^{[d]}(X)$, respectively. For $\mathcal{G}^{[d]}$ and $\mathcal{F}^{[d]}$, we use similar notations to that used for $X^{[d]}$: namely, an element of either of these groups is written as $S = (S_\epsilon : \epsilon \in \{0,1\}^d)$. In particular, $\mathcal{F}^{[d]} = \{S \in \mathcal{G}^{[d]} : S_\emptyset = \mathrm{id}\}$.

For convenience, we denote the orbit closure of $\mathbf{x} \in X^{[d]}$ under $\mathcal{F}^{[d]}$ by $\overline{\mathcal{F}^{[d]}}(\mathbf{x})$, instead of $\overline{\mathcal{O}(\mathbf{x}, \mathcal{F}^{[d]})}$. It is easy to verify that $\mathbf{Q}^{[d]}$ is the closure in $X^{[d]}$ of

$$\{Sx^{[d]} : S \in \mathcal{F}^{[d]}, x \in X\}.$$

If x is a transitive point of X, then $\mathbf{Q}^{[d]}$ is the closed orbit of $x^{[d]}$ under the group $\mathcal{G}^{[d]}$.

2.5. Measure $\mu^{[k]}$.

2.5.1. *Notation.* When f_ϵ, $\epsilon \in V_k = \{0,1\}^d$, are 2^k real or complex valued functions on the set X, we define a function $\bigotimes_{\epsilon\in V_k} f_\epsilon$ on $X^{[k]}$ by

$$\bigotimes_{\epsilon\in V_k} f_\epsilon(\mathbf{x}) = \prod_{\epsilon\in V_k} f_\epsilon(x_\epsilon).$$

2.5.2. We define by induction a $T^{[k]}$-invariant measure $\mu^{[k]}$ on $X^{[k]}$ for every integer $k \geq 0$.

Set $X^{[0]} = X$, $T^{[0]} = T$ and $\mu^{[0]} = \mu$. Assume that $\mu^{[k]}$ is defined. Let $\mathcal{I}^{[k]}$ denote the $T^{[k]}$-invariant σ-algebra of $(X^{[k]}, \mu^{[k]}, T^{[k]})$. Identifying $X^{[k+1]}$ with $X^{[k]} \times X^{[k]}$ as explained above, we define the system $(X^{[k+1]}, \mu^{[k+1]}, T^{[k+1]})$ to be the relatively independent joining of two copies of $(X^{[k]}, \mu^{[k]}, T^{[k]})$ over $\mathcal{I}^{[k]}$. That is,

$$\mathcal{I}^{[k]} = \{A \subset X^{[k]} : T^{[k]}A = A\},$$

and

$$\mu^{[k+1]} = \mu^{[k]} \underset{\mathcal{I}^{[k]}}{\times} \mu^{[k]}.$$

Equivalently, for all bounded function $f_\epsilon, \epsilon \in V_{k+1}$ of X,

$$\int_{X^{[k+1]}} \bigotimes_{\epsilon\in V_{k+1}} f_\epsilon \, d\mu^{[k+1]} = \int_{X^{[k]}} \mathbb{E}\Big(\bigotimes_{\eta\in V_k} f_{\eta 0}\Big|\mathcal{I}^{[k]}\Big)\mathbb{E}\Big(\bigotimes_{\eta\in V_k} f_{\eta 1}\Big|\mathcal{I}^{[k]}\Big) \, d\mu^{[k]}. \tag{2.4}$$

Since (X,μ,T) is ergodic, $\mathcal{I}^{[0]}$ is the trivial σ-algebra and $\mu^{[1]}=\mu\times\mu$. If (X,μ,T) is weakly mixing, then by induction $\mathcal{I}^{[k]}$ is trivial and $\mu^{[k]}$ is the 2^k Cartesian power $\mu^{\bigotimes 2^k}$ of μ for $k\geq 1$.

We now give an equivalent formulation of the definition of these measures. For an integer $k\geq 1$, let (Ω_k,P_k) be the system corresponding to the σ-algebra $\mathcal{I}^{[k]}$ and let

$$\mu^{[k]}=\int_{\Omega_k}\mu^{[k]}_\omega\ dP_k(\omega) \tag{2.5}$$

denote the ergodic decomposition of $\mu^{[k]}$ under $T^{[k]}$. Then by definition

$$\mu^{[k+1]}=\int_{\Omega_k}\mu^{[k]}_\omega\times\mu^{[k]}_\omega\ dP_k(\omega). \tag{2.6}$$

We generalize this formula. For $k,l\geq 1$, the concatenation of an element α of V_k with an element β of V_l is the element $\alpha\beta$ of V_{k+l}. This defines a bijection of $V_k\times V_l$ onto V_{k+l} and gives the identification $(X^{[k]})^{[l]}=X^{[k+1]}$. By [9, Lemma 3.1.]

$$\mu^{[k+l]}=\int_{\Omega_k}(\mu^{[k]}_\omega)^{[l]}\ dP_k(\omega). \tag{2.7}$$

2.6. **Characteristic factors** (Z_k,μ_k).

2.6.1. Notice that in [9], $\mathcal{G}^k$ and $\mathcal{F}^{[k]}$ are denoted by $\mathcal{T}^{[k]}_{k-1}$ and $\mathcal{T}^{[k]}_*$ respectively. Let $\mathcal{J}^{[k]}$ denote the σ-algebra of sets on $X^{[k]}$ that are invariant under the group $\mathcal{F}^{[k]}$. On $(X^{[k]},\mu^{[k]})$, the σ-algebra $\mathcal{J}^{[k]}$ coincides with the σ-algebra of sets depending only on the coordinate $\mathbf{0}$ ([9, Proposition 3.4]).

Proposition 2.9. [9] *For all $k\in\mathbb{N}$, $(X^{[k]},\mu^{[k]})$ is ergodic for the group of side transformations $\mathcal{G}^{[d]}$. And (Ω_k,P_k) is ergodic under the action of the group $\mathcal{F}^{[k]}$.*

We consider the 2^k-1-dimensional marginals of $\mu^{[k]}$. Recall that $V_k^*=V_k\setminus\{\mathbf{0}\}$. Consider a point $\mathbf{x}\in X^{[k]}$ as a pair $(x_{\mathbf{0}},\mathbf{x}_*)$, with $x_{\mathbf{0}}\in X$ and $\mathbf{x}_*\in X^{[k]}_*$. Let $\mu^{[k]}_*$ denote the measure on $X^{[k]}_*$, which is the image of $\mu^{[k]}$ under the natural projection $\mathbf{x}\mapsto\mathbf{x}_*$ from $X^{[k]}$ onto $X^{[k]}_*$.

All the transformations belonging to $\mathcal{G}^{[k]}$ factor through the projection $X^{[k]}\to X^{[k]}_*$ and induce transformations of $X^{[k]}_*$ preserving $\mu^{[k]}_*$. This defines a measure-preserving action of the group $\mathcal{G}^{[k]}$ and of its subgroup $\mathcal{F}^{[k]}$ on $X^{[k]}_*$. The measure $\mu^{[k]}_*$ is ergodic for the action of $\mathcal{G}^{[k]}$.

On the other hand, all the transformations belonging to $\mathcal{G}^{[k]}$ factor through the projection $\mathbf{x}\mapsto x_{\mathbf{0}}$ from $X^{[k]}$ to X, and induce measure-preserving transformations of X. The transformation $T^{[k]}$ induces the transformation T on X, and each transformation belonging to $\mathcal{F}^{[k]}$ induces the trivial transformation on X. This defines a measure-preserving ergodic action of the group $\mathcal{G}^{[k]}$ on X, with a trivial restriction to the subgroup $\mathcal{F}^{[k]}$.

2.6.2. *A system of order k.* Let $\mathcal{J}_*^{[k]}$ denote the σ-algebra of subsets of $X_*^{[k]}$ which are invariant under the action of $\mathcal{F}^{[k]}$. Since the σ-algebra $\mathcal{J}^{[k]}$ coincides with the σ-algebra of sets depending only on the coordinate $\mathbf{0}$ ([9, Proposition 3.4]). Hence there exists a σ-algebra $\mathcal{Z}_{k-1}$ of X such that $\mathcal{Z}_{k-1}$ is isomorphic to $\mathcal{J}_*^{[k]}$. To be precise, for each $A \in \mathcal{J}_*^{[k]}$, there is unique $B \in \mathcal{Z}_{k-1}$ such that $1_B(x_{\mathbf{0}}) = 1_A(\mathbf{x}_*)$ for $\mu^{[k]}$-almost every $\mathbf{x} = (x_{\mathbf{0}}, \mathbf{x}_*) \in X^{[k]}$.

Definition 2.10. The σ-algebra $\mathcal{Z}_k$ is invariant under T and so defines a factor of (X, μ, T) written $(Z_k(X), \mu_k, T)$, or simply (Z_k, μ_k, T). The factor map $X \to Z_k$ is written by π_k.

$(Z_k, \mathcal{Z}_k, \mu_k, T)$ is called a *system of order k*.

$(Z_k, \mathcal{Z}_k, \mu_k)$ has a very nice structure:

Theorem 2.11. [9] *Let $(X, \mathcal{X}, \mu, T)$ be an ergodic system and $k \in \mathbb{N}$. Then the system $(Z_k, \mathcal{Z}_k, \mu_k, T)$ is a (measure theoretic) inverse limit of k-step nilsystems.*

Remark 2.12. In this section we follows from the treatment of Host and Kra. Ziegler has a different approach, see [22]. For more details about the difference between these two methods, see Leibman's notes in the appendix in [3].

2.6.3. *Properties about Z_k.* The following properties may be useful in the proof of Theorem A.

Theorem 2.13. [9, 10] *Let $k \geq 2$ is an integer and (X, T, μ) is an ergodic $(k-1)$-step nilsystem.*

1. *The measure $\mu^{[k]}$ is an invariant measure of $\mathbf{Q}^{[k]}$. $(\mathbf{Q}^{[k]}, \mu^{[k]}, \mathcal{G}^{[k]})$ is strictly ergodic.*
2. *For every $x \in X$, let $W_{k,x} = \{\mathbf{x} \in \mathbf{Q}^{[k]} : x_{\mathbf{0}} = x\}$. Then $W_{k,x} = \overline{\mathcal{F}^{[k]}}(x^{[k]})$ and it is uniquely ergodic under $\mathcal{F}^{[k]}$.*
3. *For every $x \in X$, let $\rho_{k,x}$ be the invariant measure of $W_{k,x}$. Then for every $x \in X$, $\rho_{k,Tx}$ is the image of $\rho_{k,x}$ under the translation by $T^{[k]} = (T, T, \ldots, T)$.*

3. $\mathcal{F}^{[d]}$-($\mathcal{G}^{[d]}$-)strictly ergodic models and pointwise convergence

In this section we explain how to give the models in Theorem A, and leave the proof of the Theorem A in the next section. Also we will prove Theorem B in this section.

3.1. $\mathcal{F}^{[d]}$-strictly ergodic model and $\mathcal{G}^{[d]}$-strictly ergodic model.

Definition 3.1. Let $(X, \mathcal{X}, \mu, T)$ be an ergodic dynamical system and $(\hat{X}, \hat{T})$ be its model. For $d \in \mathbb{N}$, $(\hat{X}, \hat{T})$ is called an $\mathcal{F}^{[d]}$*-strictly ergodic model* for $(X, \mathcal{X}, \mu, T)$ if $(\hat{X}, \hat{T})$ is a strictly ergodic model and $(\overline{\mathcal{F}^{[d]}}(x^{[d]}), \mathcal{F}^{[d]})$ is uniquely ergodic for all $x \in \hat{X}$.

For $d \in \mathbb{N}$, $(\hat{X}, \hat{T})$ is called an $\mathcal{G}^{[d]}$*-strictly ergodic model* for $(X, \mathcal{X}, \mu, T)$ if $(\hat{X}, \hat{T})$ is a strictly ergodic model and $(\mathbf{Q}^{[d]}, \mathcal{G}^{[d]})$ is uniquely ergodic.

Remark 3.2. Notice that not every uniquely ergodic system has $\mathcal{F}^{[d]}$-strictly ergodic model. For example, let $(X, \mathcal{X}, \mu, T)$ be a Kronecker system. By Theorem 2.4, we may assume that (X, T) is a topologically weakly mixing minimal system and it is strictly ergodic. By [18, Theorem 3.11.] $(\overline{\mathcal{F}^{[d]}}(x^{[d]}), \mathcal{F}^{[d]})$ is minimal for all $x \in X$ and $\overline{\mathcal{F}^{[d]}}(x^{[d]}) = \{x\} \times X_*^{[d]}$. It is easy to see that $\delta_x \times \mu^{\bigotimes 2^d-1}$ and $\mu_*^{[d]}$ are two different invariant measures on it.

3.2. **Construction of models.**

3.2.1. By above definitions we restate Theorem A as follows:

Theorem 3.3. *Every ergodic dynamical system has an $\mathcal{F}^{[d]}$ and $\mathcal{G}^{[d]}$ strictly ergodic model (X, T) for all $d \in \mathbb{N}$.*

It is easy to see that Theorem 3.3 holds for Kronecker system since for group rotations the Harr measure is the unique invariant measure. In Section 4.2, we show it holds for weakly mixing system. After that we show the theorem by induction on d. Now we show idea of the proof. Let $\pi_d : X \to Z_d$ be the factor map from X to its d-step nilfactor Z_d. By definition, Z_d may be regarded as a topological system in the natural way. By Theorem 2.5, there is a uniquely ergodic model $(\hat{X}, \hat{\mathcal{X}}, \hat{\mu}, T)$ for $(X, \mathcal{X}, \mu, T)$ and a factor map $\hat{\pi}_d : \hat{X} \to Z_d$ which is a model for $\pi_d : X \to Z_d$.

$$\begin{array}{ccc} X & \xrightarrow{\phi} & \hat{X} \\ {\scriptstyle \pi_d}\downarrow & & \downarrow{\scriptstyle \hat{\pi}_d} \\ Z_d & \longrightarrow & Z_d \end{array}$$

The difficult part is to verify that $(\hat{X}, T)$ is what we need.

3.3. ***d*-step almost automorphic systems.** d-step almost automorphic systems are defined and studied in [12] which are the generalization of Veech's almost automorphic systems.

Definition 3.4. Let (X, T) be a minimal t.d.s. and $d \in \mathbb{N}$. (X, T) is called a *d-step almost automorphic* system if it is an almost one-to-one extension of a d-step nilsystem.

See [12] for more discussion about d-step almost automorphy. In this subsection we will show that in Theorem A we can also require the models are d-step almost automorphic systems. To do so, first we state Furstenberg-Weiss's almost one-to-one Theorem.

Theorem 3.5 (Furstenberg-Weiss)**.** [6] *Let (Y, T) be a non-periodic minimal t.d.s., and let $\pi' : X' \to Y$ be an extension of (Y, T) with (X', T) topologically transitive and X' a compact metric space.*

$$\begin{array}{ccc} X' & \xrightarrow{\theta} & X \\ {\scriptstyle \pi'}\downarrow & & \downarrow{\scriptstyle \pi} \\ Y & \longrightarrow & Y \end{array}$$

Then there exists an almost 1-1 minimal extension $\pi : (X, T) \to (Y, T)$, a Borel subset $X'_0 \subseteq X'$ and a Borel measurable map $\theta : X'_0 \to X$ satisfying:

1. *$\theta \circ T = T \circ \theta$;*
2. *$\pi \circ \theta = \pi'$;*
3. *θ is a Borel isomorphism of X'_0 onto its image $X_0 = \theta(X'_0) \subseteq X$;*
4. *$\mu(X'_0) = 1$ for any T-invariant measure μ on X'.*
5. *if (X', T) is uniquely ergodic, then (X, T) can be chosen to be uniquely (hence strictly) ergodic.*

Remark 3.6. In [6, Theorem 1], (1)-(4) are stated. From the proof of the theorem given in [6], we have (5), which is pointed out in [8].

Let $(X, \mathcal{X}, \mu, T)$ be an ergodic system with non-trivial nil-factors (non-triviality here means infinity) and $d \in \mathbb{N}$. Let $\pi_d : X \to Z_d$ be the factor map from X to its d-step nilfactor Z_d. By definition, Z_d may be regarded as a t.d.s. in the natural way. By Weiss's theorem [19], there is a uniquely ergodic model $(\hat{X}', \hat{\mathcal{X}}', \hat{\mu}, T)$ for $(X, \mathcal{X}, \mu, T)$ and a factor map $\hat{\pi}'_d : \hat{X}' \to Z_d$ which is a model for $\pi_d : X \to Z_d$.

$$\begin{array}{ccccc} X & \xrightarrow{\phi} & \hat{X}' & \xrightarrow{\theta} & \hat{X} \\ {\scriptstyle \pi_d}\downarrow & & \downarrow{\scriptstyle \hat{\pi}'_d} & & \downarrow{\scriptstyle \hat{\pi}_d} \\ Z_d & \longrightarrow & Z_d & \longrightarrow & Z_d \end{array}$$

Now by Theorem 3.5, $\hat{\pi}'_d : \hat{X}' \to Z_d$ may be replaced by $\hat{\pi}_d : \hat{X} \to Z_d$, where $\hat{\pi}_d$ is almost 1-1 and $\hat{X}'$ and $\hat{X}$ are measure theoretically isomorphic. In particular, $(\hat{X}, T)$ is a strictly ergodic model for $(X, \mathcal{X}, \mu, T)$.

As we described in the introduction, one once we have a model $\hat{\pi} : \hat{X} \longrightarrow Z_d$ then it is $\mathcal{F}^{[d]}$ and $\mathcal{G}^{[d]}$ models. Hence combining above discussion with Theorem A, we have

Theorem 3.7. *Let $d \in \mathbb{N}$. Then every ergodic m.p.t. with a non-trivial d-step nilfactor has an $\mathcal{F}^{[d]}$ and $\mathcal{G}^{[d]}$ strictly ergodic model (X, T) which is a d-step almost automorphic system.*

3.4. Pointwise convergence along cubes.

The following equation is easy to be verified.

Lemma 3.8. *Let $\{a_i\}, \{b_i\} \subseteq \mathbb{C}$. Then*

$$\prod_{i=1}^{k} a_i - \prod_{i=1}^{k} b_i = (a_1 - b_1) b_2 \dots b_k + a_1 (a_2 - b_2) b_3 \dots b_k + a_1 \dots a_{k-1}(a_k - b_k). \tag{3.1}$$

Proof of Theorem B. (1) Since $(X, \mathcal{X}, \mu, T)$ has an $\mathcal{F}^{[d]}$-strictly ergodic model, we may assume that (X, T) itself is a topological minimal system and μ is its unique measure such that $(\overline{\mathcal{F}^{[d]}(x^{[d]})}, \mathcal{F}^{[d]})$ is uniquely ergodic for all $x \in X$.

Without loss of generality, we assume that for all $\emptyset \neq \epsilon \subset [d]$, $\|f_\epsilon\|_\infty \leq 1$. Let $\delta > 0$, and choose continuous function g_ϵ such that $\|g_\epsilon\|_\infty \leq 1$ and $\|f_\epsilon - g_\epsilon\|_1 < \delta/2^d$

for all $\emptyset \neq \epsilon \subset [d]$. By Lemma 3.8, we have

$$\begin{aligned}
&\left| \frac{1}{N^d} \sum_{\mathbf{n}\in[0,N-1]^d} \prod_{\emptyset\neq\epsilon\subset[d]} f_\epsilon(T^{\mathbf{n}\cdot\epsilon}x) - \frac{1}{N^d} \sum_{\mathbf{n}\in[0,N-1]^d} \prod_{\emptyset\neq\epsilon\subset[d]} g_\epsilon(T^{\mathbf{n}\cdot\epsilon}x) \right| \\
=& \left| \frac{1}{N^d} \sum_{\mathbf{n}\in[0,N-1]^d} \Big[\prod_{\emptyset\neq\epsilon\subset[d]} f_\epsilon(T^{\mathbf{n}\cdot\epsilon}x) - \prod_{\emptyset\neq\epsilon\subset[d]} g_\epsilon(T^{\mathbf{n}\cdot\epsilon}x) \Big] \right| \\
\le& \sum_{\emptyset\neq\epsilon\subset[d]} \left[\frac{1}{N^d} \sum_{\mathbf{n}\in[0,N-1]^d} \Big| f_\epsilon(T^{\mathbf{n}\cdot\epsilon}x) - g_\epsilon(T^{\mathbf{n}\cdot\epsilon}x) \Big| \right].
\end{aligned}$$

Now by Pointwise Ergodic Theorem for $\mathbb{Z}^d$ [17] we have that for all $\emptyset \neq \epsilon \subset [d]$

$$\frac{1}{N^d} \sum_{\mathbf{n}\in[0,N-1]^d} \Big| f_\epsilon(T^{\mathbf{n}\cdot\epsilon}x) - g_\epsilon(T^{\mathbf{n}\cdot\epsilon}x) \Big| \longrightarrow \|f_\epsilon - g_\epsilon\|_1, \quad N\to\infty,\ \mu\ a.e.$$

Hence

$$\begin{aligned}
&\left| \frac{1}{N^d} \sum_{\mathbf{n}\in[0,N-1]^d} \prod_{\emptyset\neq\epsilon\subset[d]} f_\epsilon(T^{\mathbf{n}\cdot\epsilon}x) - \frac{1}{N^d} \sum_{\mathbf{n}\in[0,N-1]^d} \prod_{\emptyset\neq\epsilon\subset[d]} g_\epsilon(T^{\mathbf{n}\cdot\epsilon}x) \right| \\
\le& \sum_{\emptyset\neq\epsilon\subset[d]} \Big[\frac{1}{N^d} \sum_{\mathbf{n}\in[0,N-1]^d} \Big| f_\epsilon(T^{\mathbf{n}\cdot\epsilon}x) - g_\epsilon(T^{\mathbf{n}\cdot\epsilon}x) \Big| \Big] \qquad (3.2) \\
\longrightarrow& \sum_{\emptyset\neq\epsilon\subset[d]} \|f_\epsilon - g_\epsilon\|_1 \le \delta, \quad N\to\infty,\ \mu\ a.e.
\end{aligned}$$

Note that

$$\begin{aligned}
&\frac{1}{N^d} \sum_{\mathbf{n}\in[0,N-1]^d} \prod_{\emptyset\neq\epsilon\subset[d]} g_\epsilon(T^{\mathbf{n}\cdot\epsilon}x) = \frac{1}{N^d} \sum_{\substack{0\le n_1\le N-1,\\ \dots,\\ 1\le n_d\le N-1}} \prod_{\epsilon\in V_d^*} g_\epsilon(T^{\epsilon_1 n_1+\ldots+\epsilon_d n_d}x) \\
=& \frac{1}{N^d} \sum_{\substack{0\le n_1\le N-1,\\ \dots,\\ 1\le n_d\le N-1}} \bigotimes_{\epsilon\in V_d^*} g_\epsilon\Big((T_1^{[d]})^{n_1} \dots (T_d^{[d]})^{n_d} x^{[d]} \Big).
\end{aligned}$$

Since $\bigotimes_{\epsilon\in V_k^*} g_\epsilon : X_*^{[d]} \to \mathbb{R}$ is continuous and $(\overline{\mathcal{F}^{[d]}}(x^{[d]}), \mathcal{F}^{[d]})$ is uniquely ergodic, by Theorem 2.2, $\frac{1}{N^d} \sum_{\mathbf{n}\in[0,N-1]^d} \prod_{\emptyset\neq\epsilon\subset[d]} g_\epsilon(T^{\mathbf{n}\cdot\epsilon}x)$ converges pointwisely. Together with (3.2),

$$\frac{1}{N^d} \sum_{\mathbf{n}\in[0,N-1]^d} \prod_{\emptyset\neq\epsilon\subset[d]} f_\epsilon(T^{\mathbf{n}\cdot\epsilon}x)$$

converge μ a.e..

Similarly, one has (2). The proof is completed. □

Remark 3.9. It is easy to see that if (1.1) holds for all d, then we have (1.2) holds for all d. That is, (1.1) is more fundamental. For example, if we want to get $\mathcal{G}^{[1]}$-case:

$$\frac{1}{N^2}\sum_{0\le n_1,n_2\le N-1} f_0(T^{n_1}x)f_1(T^{n_1+n_2}x),$$

then what need do is in the $\mathcal{F}^{[2]}$-case

$$\frac{1}{N^2}\sum_{0\le n_1,n_2\le N-1} f_{01}(T^{n_1}x)f_{10}(T^{n_2}x)f_{11}(T^{n_1+n_2}x)$$

by setting $f_{00}=f_0, f_{10}=1$ and $f_{11}=f_1$.

4. Proof of Theorem A

In this section we give a proof for Theorem A. To make the idea of the proof clearer before going into the proof for the general case we show the cases when $d=1$ and $d=2$ first. We also give a proof for weakly mixing systems for independent interest. Finally we show the general case by induction.

4.1. **Case when $d=1$.** By Jewett-Krieger's Theorem, every ergodic system has a strictly ergodic model. Now we show this model is $\mathcal{F}^{[1]}$-strictly ergodic. Let (X,T) be a strictly ergodic system and let μ be its unique T-invariant measure. Note that $\mathcal{F}^{[1]}=\langle \mathrm{id}\times T\rangle$. Hence for all $x\in X$,

$$\overline{\mathcal{F}^{[1]}}(x^{[1]})=\{x\}\times X.$$

Since (X,T) is uniquely ergodic, $\delta_x\times\mu$ is the only $\mathcal{F}^{[1]}$-invariant measure of $\overline{\mathcal{F}^{[1]}}(x^{[1]})$. In this case Theorem A(1) is nothing but Birkhorff pointwise ergodic theorem.

Now consider $\mathbf{Q}^{[1]}$. Since $\mathcal{G}^{[1]}=\langle T\times T, \mathrm{id}\times T\rangle$, it is easy to see that $\mathbf{Q}^{[1]}=X\times X$. Let λ be a $\mathcal{G}^{[1]}$-invariant measure of $(X^{[1]},\mathcal{X}^{[1]})=(X\times X,\mathcal{X}\times\mathcal{X})$. Since λ is $T\times T$-invariant, it is a self-joining of $(X,\mathcal{X},\mu,T)$ and has μ as its marginal. Let

$$\lambda=\int_X \delta_x\times\lambda_x\, d\mu(x) \tag{4.1}$$

be the disintegration of λ over μ. Since λ is $\mathrm{id}\times T$-invariant, we have

$$\lambda=\mathrm{id}\times T\lambda=\int_X \delta_x\times T\lambda_x\, d\mu(x).$$

The uniqueness of disintegration implies that

$$T\lambda_x=\lambda_x, \mu\ a.e.$$

Since $(X,\mathcal{X},T)$ is uniquely ergodic, $\lambda_x=\mu,\ \mu$ a.e. Thus by (4.1) one has that

$$\lambda=\int_X \delta_x\times\lambda_x\, d\mu(x)=\int_X \delta_x\times\mu\, d\mu(x)=\mu\times\mu.$$

Hence $(\mathbf{Q}^{[1]},\mathcal{G}^{[1]})$ is uniquely ergodic, and $\mu^{[1]}=\mu\times\mu$ is its unique $\mathcal{G}^{[1]}$-invariant measure.

4.2. **Weakly mixing systems.** In this subsection we show Theorem A holds for weakly mixing systems. This result relies on the following proposition.

Proposition 4.1. *Let (X,T) be uniquely ergodic, $(X,\mathcal{X},\mu,T)$ be weakly mixing and $d\in\mathbb{N}$. Then*

(1) *$(X^{[d]},\mathcal{G}^{[d]})$ is uniquely ergodic with the unique measure $\mu^{[d]}=\underbrace{\mu\times\ldots\times\mu}_{2^d\ times}$.*

(2) *$(X_*^{[d]},\mathcal{F}^{[d]})$ is uniquely ergodic with the unique measure $\mu_*^{[d]}=\underbrace{\mu\times\ldots\times\mu}_{2^d-1\ times}$.*

Proof. We prove the result inductively. First we show the case when $d=1$. In this case $\mathcal{F}^{[1]}=\langle \mathrm{id}\times T\rangle$ and $\mathcal{G}^{[1]}=\langle \mathrm{id}\times T, T\times T\rangle$. Hence $(X_*^{[1]},\mathcal{X}_*^{[1]},\mathcal{F}_*^{[1]})=(X,\mathcal{X},T)$, and it follows that $\mu_*^{[1]}=\mu$ is the unique T-invariant measure. Let λ be a $\mathcal{G}^{[1]}$-invariant measure of $(X^{[1]},\mathcal{X}^{[1]})=(X\times X,\mathcal{X}\times\mathcal{X})$. By the argument in subsection 4.1, we know that $\lambda=\mu^{[1]}=\mu\times\mu$.

Now assume the statements hold for $d-1$, and we show the case for d. Let λ be a $\mathcal{G}^{[d]}$-invariant measure of $(X^{[d]},\mathcal{X}^{[d]})$. Let

$$p_1:(X^{[d]},\mathcal{G}^{[d]})\to(X^{[d-1]},\mathcal{G}^{[d-1]});\ \mathbf{x}=(\mathbf{x}',\mathbf{x}'')\mapsto\mathbf{x}'$$
$$p_2:(X^{[d]},\mathcal{G}^{[d]})\to(X^{[d-1]},\mathcal{G}^{[d-1]});\ \mathbf{x}=(\mathbf{x}',\mathbf{x}'')\mapsto\mathbf{x}''$$

be the projections. Then $(p_2)_*(\lambda)$ is a $\mathcal{G}^{[d-1]}$-invariant measure of $X^{[d-1]}$. By inductive assumption, $(p_2)_*(\lambda)=\mu^{[d-1]}$. Let

$$\lambda=\int_{X^{[d-1]}}\lambda_{\mathbf{x}}\times\delta_{\mathbf{x}}\ d\mu^{[d-1]}(\mathbf{x}) \tag{4.2}$$

be the disintegration of λ over $\mu^{[d-1]}$. Since λ is $T_d^{[d]}=\mathrm{id}^{[d-1]}\times T^{[d-1]}$-invariant, we have

$$\begin{aligned}\lambda &= \mathrm{id}^{[d-1]}\times T^{[d-1]}\lambda=\int_{X^{[d-1]}}\lambda_{\mathbf{x}}\times T^{[d-1]}\delta_{\mathbf{x}}\ d\mu^{[d-1]}(\mathbf{x})\\ &= \int_{X^{[d-1]}}\lambda_{\mathbf{x}}\times\delta_{T^{[d-1]}\mathbf{x}}\ d\mu^{[d-1]}(\mathbf{x})\\ &= \int_{X^{[d-1]}}\lambda_{(T^{[d-1]})^{-1}\mathbf{x}}\times\delta_{\mathbf{x}}\ d\mu^{[d-1]}(\mathbf{x}).\end{aligned}$$

The uniqueness of disintegration implies that

$$\lambda_{(T^{[d-1]})^{-1}\mathbf{x}}=\lambda_{\mathbf{x}},\quad \mu^{[d-1]}\ a.e.\ \mathbf{x}\in X^{[d-1]}. \tag{4.3}$$

Define

$$F:(X^{[d-1]},\mathcal{X}^{[d-1]},T^{[d-1]})\longrightarrow M(X^{[d-1]}):\ \mathbf{x}\mapsto\lambda_{\mathbf{x}}.$$

By (4.3), F is a $T^{[d-1]}$-invariant $M(X^{[d-1]})$-value function. Since $(X,\mathcal{X},\mu,T)$ is weakly mixing, $(X^{[d-1]},\mathcal{X}^{[d-1]},T^{[d-1]})$ is ergodic and hence $\lambda_{\mathbf{x}}=\nu,\ \mu^{[d-1]}$ a.e. for some $\nu\in M(X^{[d-1]})$. Thus by (4.2) one has that

$$\lambda=\int_{X^{[d-1]}}\lambda_{\mathbf{x}}\times\delta_{\mathbf{x}}\ d\mu^{[d-1]}(\mathbf{x})=\int_{X^{[d-1]}}\nu\times\delta_{\mathbf{x}}\ d\mu^{[d-1]}(\mathbf{x})=\nu\times\mu^{[d-1]}.$$

Then we have that $\nu = (p_1)_*(\lambda)$ is a $\mathcal{G}^{[d-1]}$-invariant measure of $X^{[d-1]}$. By inductive assumption, $\mu^{[d-1]}$ is the only $\mathcal{G}^{[d-1]}$-invariant measure of $X^{[d-1]}$ and hence $\nu = (p_1)_*(\lambda) = \mu^{[d-1]}$. Thus $\lambda = \mu^{[d-1]} \times \mu^{[d-1]} = \mu^{[d]}$. That is, $(X^{[d]}, \mathcal{X}^{[d]}, \mu^{[d]}, \mathcal{G}^{[d]})$ is uniquely ergodic.

Now we show that $(X_*^{[d]}, \mathcal{X}_*^{[d]}, \mu_*^{[d]}, \mathcal{F}^{[d]})$ is uniquely ergodic. The proof is similar. Let λ be a $\mathcal{F}^{[d]}$-invariant measure of $(X_*^{[d]}, \mathcal{X}_*^{[d]})$. Let

$$q_1 : (X_*^{[d]}, \mathcal{F}^{[d]}) \to (X_*^{[d-1]}, \mathcal{F}^{[d-1]});\ \mathbf{x} = (\mathbf{x}'_*, \mathbf{x}'') \mapsto \mathbf{x}'_*$$
$$q_2 : (X^{[d]}, \mathcal{F}^{[d]}) \to (X^{[d-1]}, \mathcal{G}^{[d-1]});\ \mathbf{x} = (\mathbf{x}'_*, \mathbf{x}'') \mapsto \mathbf{x}''$$

be the projections. Then $(q_2)_*(\lambda)$ is a $\mathcal{G}^{[d-1]}$-invariant measure of $X^{[d-1]}$. By inductive assumption, $(q_2)_*(\lambda) = \mu^{[d-1]}$. Let

$$\lambda = \int_{X^{[d-1]}} \lambda_{\mathbf{x}} \times \delta_{\mathbf{x}}\ d\mu^{[d-1]}(\mathbf{x}) \tag{4.4}$$

be the disintegration of λ over $\mu^{[d-1]}$. Since λ is $T_d^{[d]} = \mathrm{id}^{[d-1]} \times T^{[d-1]}$-invariant, we have

$$\begin{aligned} \lambda &= \mathrm{id}^{[d-1]} \times T^{[d-1]}\lambda = \int_{X^{[d-1]}} \lambda_{\mathbf{x}} \times T^{[d-1]}\delta_{\mathbf{x}}\ d\mu^{[d-1]}(\mathbf{x}) \\ &= \int_{X^{[d-1]}} \lambda_{\mathbf{x}} \times \delta_{T^{[d-1]}\mathbf{x}}\ d\mu^{[d-1]}(\mathbf{x}) \\ &= \int_{X^{[d-1]}} \lambda_{(T^{[d-1]})^{-1}\mathbf{x}} \times \delta_{\mathbf{x}}\ d\mu^{[d-1]}(\mathbf{x}). \end{aligned}$$

The uniqueness of disintegration implies that

$$\lambda_{(T^{[d-1]})^{-1}\mathbf{x}} = \lambda_{\mathbf{x}}, \quad \mu^{[d-1]}\ a.e. \tag{4.5}$$

Define

$$F : (X^{[d-1]}, \mathcal{X}^{[d-1]}, T^{[d-1]}) \longrightarrow M(X_*^{[d-1]}) :\ \mathbf{x} \mapsto \lambda_{\mathbf{x}}.$$

By (4.5), F is a $T^{[d-1]}$-invariant $M(X_*^{[d-1]})$-value function. Since $(X, \mathcal{X}, \mu, T)$ is weakly mixing, $(X^{[d-1]}, \mathcal{X}^{[d-1]}, T^{[d-1]})$ is ergodic and hence $\lambda_{\mathbf{x}} = \nu$, $\mu^{[d-1]}$ a.e. for some $\nu \in M(X_*^{[d-1]})$. Thus by (4.4) one has that

$$\lambda = \int_{X^{[d-1]}} \lambda_{\mathbf{x}} \times \delta_{\mathbf{x}}\ d\mu^{[d-1]}(\mathbf{x}) = \int_{X^{[d-1]}} \nu \times \delta_{\mathbf{x}}\ d\mu^{[d-1]}(\mathbf{x}) = \nu \times \mu^{[d-1]}.$$

Then we have that $\nu = (q_1)_*(\lambda)$ is a $\mathcal{F}^{[d-1]}$-invariant measure of $X_*^{[d-1]}$. By inductive assumption, $\mu_*^{[d-1]}$ is the only $\mathcal{F}^{[d-1]}$-invariant measure of $X_*^{[d-1]}$ and $\nu = (q_1)_*(\lambda) = \mu_*^{[d-1]}$. Thus $\lambda = \mu_*^{[d-1]} \times \mu^{[d-1]} = \mu_*^{[d]}$. Hence $(X_*^{[d]}, \mathcal{X}_*^{[d]}, \mu_*^{[d]}, \mathcal{F}^{[d]})$ is uniquely ergodic. The proof is completed. □

Theorem 4.2. *If $(X, \mathcal{X}, \mu, T)$ is a weakly mixing m.p.t., then it has an $\mathcal{F}^{[d]}$ and $\mathcal{G}^{[d]}$ strictly ergodic model for all $d \in \mathbb{N}$.*

Proof. By Jewett-Krieger' Theorem, $(X, \mathcal{X}, \mu, T)$ has a uniquely ergodic model. Without loss of generality, we assume that (X, T) itself is a minimal t.d.s. and μ is its unique T-invariant measure. By [18, Theorem 3.11.], $(\mathbf{Q}^{[d]} = X^{[d]}, \mathcal{G}^{[d]})$ is minimal, and for all $x \in X$, $(\overline{\mathcal{F}^{[d]}}(x^{[d]}), \mathcal{F}^{[d]})$ is minimal and $\overline{\mathcal{F}^{[d]}}(x^{[d]}) = \{x\} \times X_*^{[d]} = \{x\} \times X^{2^d-1}$. By Proposition 4.1, $(\mathbf{Q}^{[d]}, \mathcal{G}^{[d]})$ and $(\overline{\mathcal{F}^{[d]}}(x^{[d]}), \mathcal{F}^{[d]})$ (for all $x \in X$) are uniquely ergodic . Hence it has an $\mathcal{F}^{[d]}$ and $\mathcal{G}^{[d]}$ strictly ergodic model for all $d \in \mathbb{N}$. □

4.3. **Case when** $d = 2$**.** In this case we can give the explicit description of the unique measure. People familiar with the materials can read the proof for the general case directly.

4.3.1. *Graph joinings.* Let $\phi : (X, \mathcal{X}, \mu, T) \to (Y, \mathcal{Y}, \nu, T)$ be a homomorphism of ergodic systems. Let $\mathrm{id} \times \phi : X \to X \times Y, x \mapsto (x, \phi(x))$. Define

$$\mathrm{gr}(\mu, \phi) = \int_X \delta_x \times \delta_{\phi(x)} \, d\mu(x) = (\mathrm{id} \times \phi)_*(\mu). \tag{4.6}$$

It is called a *graph joining* of ϕ. Equivalently, $\mathrm{gr}(\mu, \phi)$ is defined by

$$\mathrm{gr}(\mu, \phi)(A \times B) = \mu(A \cap \phi^{-1}B), \ \forall A \in \mathcal{X}, B \in \mathcal{Y}. \tag{4.7}$$

4.3.2. *Kronecker factor* Z_1*.* The Kronecker factor of the ergodic system (X, μ, T) is an ergodic rotation and we denote it by $(Z_1(X), t_1)$, or more simply (Z_1, t_1). Let μ_1 denote the Haar measure of Z_1, and $\pi_{X,1}$ or π_1, denote the factor map $X \to Z_1$.

For $s \in Z_1$, let $\mu_{1,s}$ denote the image of the measure μ_1 under the map $z \mapsto (z, sz)$ from Z_1 to Z_1^2, i.e. $\mu_{1,s} = \mathrm{gr}(\mu_1, s)$. This measure is invariant under $T^{[1]} = T \times T$ and is a self-joining of the rotation (Z_1, t_1). Let μ_s denote the relatively independent joining of μ over $\mu_{1,s}$. This means that for bounded measurable functions f and g on X,

$$\int_{Z_1 \times Z_1} f(x_0) g(x_1) \, d\mu_s(x_0, x_1) = \int_{Z_1} \mathbb{E}(f|\mathcal{Z}_1)(z) \mathbb{E}(g|\mathcal{Z}_1)(sz) \, d\mu_1(z). \tag{4.8}$$

where we view the conditional expectations relative to $\mathcal{Z}_1$ as functions defined on Z_1.

It is a classical result that the invariant σ-algebra $\mathcal{I}^{[1]}$ of $(X \times X, \mu \times \mu, T \times T)$ consists in sets of the form

$$\{(x, y) \in X \times X : \pi_1(x) - \pi_1(y) \in A\} \tag{4.9}$$

where $A \in \mathcal{Z}_1$. Hence $\mathcal{I}^{[1]}$ is isomorphic to $\mathcal{Z}_1$. Let $\phi : (X \times X, \mathcal{X} \times \mathcal{X}) \to (\Omega_1, \mathcal{I}^{[1]}, P_1)$ be the factor map and let $\psi : (\Omega_1, \mathcal{I}^{[1]}, P_1) \to (Z_1, \mathcal{Z}_1, \mu_1)$ be the isomorphic map. Hence we have

$$\begin{array}{c} (X \times X, \mathcal{X} \times \mathcal{X}) \xrightarrow{\phi} (\Omega_1, \mathcal{I}^{[1]}, P_1) \xleftrightarrow{\psi} (Z_1, \mathcal{Z}_1, \mu_1) \\ (x, y) \longrightarrow \phi(x, y) \longleftrightarrow s = \psi(\phi(x, y)) \end{array} \tag{4.10}$$

From this, it is not difficult to deduce that the ergodic decomposition of $\mu \times \mu$ under $T \times T$ can be written as

$$\mu \times \mu = \int_{Z_1} \mu_s \, d\mu_1(s). \tag{4.11}$$

In particular, for μ_1-almost every s, the measure μ_s is ergodic for $T \times T$. For an integer $d > 0$ we have

$$\mu^{[d+1]} = \int_{Z_1} (\mu_s)^{[d]}\, d\mu_1(s). \tag{4.12}$$

Especially, we have

$$\mu^{[2]} = \int_{Z_1} \mu_s \times \mu_s\, d\mu_1(s). \tag{4.13}$$

4.3.3. $\mathcal{G}^{[2]}$-*actions.* Let $\pi_1 : X \to Z_1$ be the factor map from X to its Kronecker factor Z_1. Since Z_1 is a group rotation, it may be regarded as a topological system in the natural way. By Weiss's Theorem, there is a uniquely ergodic model $(\hat{X}, \hat{\mathcal{X}}, \hat{\mu}, T)$ for $(X, \mathcal{X}, \mu, T)$ and a factor map $\hat{\pi}_1 : \hat{X} \to Z_1$ which is a model for $\pi_1 : X \to Z_1$.

$$\begin{array}{ccc} X & \longrightarrow & \hat{X} \\ {\scriptstyle \pi_1}\big\downarrow & & \big\downarrow{\scriptstyle \hat{\pi}_1} \\ Z_1 & \longrightarrow & Z_1 \end{array}$$

Hence for simplicity, we may assume that $(\hat{X}, \hat{\mathcal{X}}, \hat{\mu}, T) = (X, \mathcal{X}, \mu, T)$ and $\pi_1 = \hat{\pi}_1$. Now we show that $(\mathbf{Q}^{[2]}, \mu^{[2]}, \mathcal{G}^{[2]})$ is uniquely ergodic.

Let λ be a $\mathcal{G}^{[2]}$-invariant measure of $\mathbf{Q}^{[2]}$. Let

$$p_1 : (\mathbf{Q}^{[2]}, \mathcal{G}^{[2]}) \to (\mathbf{Q}^{[1]}, \mathcal{G}^{[2]});\ \mathbf{x} = (\mathbf{x}', \mathbf{x}'') \mapsto \mathbf{x}'$$
$$p_2 : (\mathbf{Q}^{[2]}, \mathcal{G}^{[2]}) \to (\mathbf{Q}^{[1]}, \mathcal{G}^{[2]});\ \mathbf{x} = (\mathbf{x}', \mathbf{x}'') \mapsto \mathbf{x}''$$

be the projections. Then $(p_2)_*(\lambda)$ is a $\mathcal{G}^{[2]}$-invariant measure of $\mathbf{Q}^{[1]} = X^{[1]}$. Note that $\mathcal{G}^{[2]}$ acts on $\mathbf{Q}^{[1]}$ as $\mathcal{G}^{[1]}$ actions. By subsection 4.1, $(p_2)_*(\lambda) = \mu^{[1]} = \mu \times \mu$. Hence let

$$\lambda = \int_{X^2} \lambda_{(x,y)} \times \delta_{(x,y)}\, d\mu \times \mu(x, y) \tag{4.14}$$

be the disintegration of λ over $\mu^{[1]}$. Since λ is $T_2^{[2]} = \mathrm{id}^{[1]} \times T^{[1]}$-invariant, we have

$$\begin{aligned} \lambda &= \mathrm{id}^{[1]} \times T^{[1]}\lambda = \int_{X^2} \lambda_{(x,y)} \times T^{[1]}\delta_{(x,y)}\, d\mu \times \mu(x, y) \\ &= \int_{X^2} \lambda_{(x,y)} \times \delta_{T^{[1]}(x,y)}\, d\mu \times \mu(x, y) \\ &= \int_{X^2} \lambda_{(T^{[1]})^{-1}(x,y)} \times \delta_{(x,y)}\, d\mu \times \mu(x, y). \end{aligned}$$

The uniqueness of disintegration implies that

$$\lambda_{(T^{[1]})^{-1}(x,y)} = \lambda_{(x,y)}, \quad \mu^{[1]} = \mu \times \mu \ a.e. \tag{4.15}$$

Define

$$F : (\mathbf{Q}^{[1]} = X^{[1]}, T^{[1]}) \longrightarrow M(X^{[1]}) :\ (x, y) \mapsto \lambda_{(x,y)}.$$

By (4.15), F is a $T^{[1]}$-invariant $M(X^{[1]})$-value function. Hence F is $\mathcal{I}^{[1]}$-measurable, and hence $\lambda_{(x,y)} = \lambda_{\phi(x,y)} = \lambda_s,\ \mu^{[1]}$ a.e., where ϕ is defined in (4.10).

Thus by (4.14) one has that

$$\begin{aligned}\lambda &= \int_{X^2} \lambda_{(x,y)} \times \delta_{(x,y)}\ d\mu \times \mu(x,y) = \int_{X^2} \lambda_{\phi(x,y)} \times \delta_{(x,y)}\ d\mu \times \mu(x,y)\\ &= \int_{Z_1}\int_{X^2} \lambda_s \times \delta_{(x,y)}\ d\mu_s(x,y) d\mu_1(s)\\ &= \int_{Z_1} \lambda_s \times \Big(\int_{X^2} \delta_{(x,y)}\ d\mu_s(x,y)\Big) d\mu_1(s)\\ &= \int_{Z_1} \lambda_s \times \mu_s\ d\mu_1(s)\end{aligned}$$

Let $\pi_1^{[2]} : (\mathbf{Q}^{[2]}(X), \mathcal{G}^{[2]}) \longrightarrow (\mathbf{Q}^{[2]}(Z_1), \mathcal{G}^{[2]})$ be the natural factor map. By Theorem 2.13, $(\mathbf{Q}^{[2]}(Z_1), \mu_1^{[2]})$ is uniquely ergodic. Hence

$$\pi_{1*}^{[2]}(\lambda) = \mu_1^{[2]} = \int_{Z_1} \mu_{1,s} \times \mu_{1,s}\ d\mu_1(s).$$

So

$$(\pi_1 \times \pi_1)_*(\lambda_s) = (\pi \times \pi)_*(\mu_s) = \mu_{1,s}.$$

Note that we have that

$$(p_1)_*(\lambda) = (p_2)_*(\lambda) = \mu^{[1]} = \mu \times \mu,$$

and hence we have

$$\mu \times \mu = \int_{Z_1} \lambda_s\ d\mu_1(s) = \int_{Z_1} \mu_s\ d\mu_1(s).$$

Hence by the uniqueness of disintegration, we have that $\lambda_s = \mu_s$, μ_1 a.e.. More precisely, if $\lambda_s \neq \mu_s$, μ_1 a.e., then $\mu_1(\{s \in Z_1 : \lambda_s \neq \mu_s\}) > 0$. So there is some function $f \in C(X \times X)$ such that

$$\mu_1\Big(\{s : \lambda_s(f) > \mu_s(f)\}\Big) > 0.$$

Let $A = \{s : \lambda_s(f) > \mu_s(f)\}$. By (4.10), we can consider A as a subset of $X \times X$:

$$A = \{s : \lambda_s(f) > \mu_s(f)\} = \{(x,y) \in X \times X : \lambda_{\phi(x,y)}(f) > \mu_{\phi(x,y)}(f)\}.$$

Hence by $\mu \times \mu = \int_{Z_1} \lambda_s\ d\mu_1(s)$ we have

$$\begin{aligned}\mu \times \mu(f \cdot 1_A) &= \int_{X^2} f \cdot 1_A\ d\mu \times \mu\\ &= \int_{Z_1}\int_{X^2} f \cdot 1_A\ d\lambda_s(x,y) d\mu_1(s)\\ &= \int_{Z_1} 1_A \int_{X^2} f\ d\lambda_s(x,y)\ d\mu_1(s)\\ &= \int_A \lambda_s(f)\ d\mu_1(s)\end{aligned}$$

Similarly, by $\mu\times\mu=\int_{Z_1}\mu_s\ d\mu_1(s)$ we have

$$\mu\times\mu(f\cdot 1_A)=\int_A \mu_s(f)\ d\mu_1(s)$$

Thus

$$0=\int_A \lambda_s(f)\ d\mu_1(s)-\int_A \mu_s(f)\ d\mu_1(s)=\int_A \Big(\lambda_s(f)-\mu_s(f)\Big)\ d\mu_1(s)>0,$$

a contradiction! Hence $\lambda_s=\mu_s$, μ_1 a.e., and

$$\lambda=\int_{Z_1}\lambda_s\times\mu_s\ d\mu_1(s)=\int_{Z_1}\mu_s\times\mu_s\ d\mu_1(s)=\mu^{[2]}.$$

That is, $(\mathbf{Q}^{[2]},\mu^{[2]},\mathcal{G}^{[2]})$ is uniquely ergodic. The proof is completed.

4.3.4. $\mathcal{F}^{[2]}$-*actions.* We use the same model as in the proof of Proposition 4.3.3.

Let λ be a $\mathcal{F}^{[2]}$-invariant measure of $\overline{\mathcal{F}^{[2]}}(x^{[2]})$. Let

$$p_1:(\overline{\mathcal{F}^{[2]}}(x^{[2]}),\mathcal{F}^{[2]})\to(\overline{\mathcal{F}^{[1]}}(x^{[1]}),\mathcal{F}^{[2]});\ \mathbf{x}=(\mathbf{x}',\mathbf{x}'')\mapsto\mathbf{x}'$$
$$p_2:(\overline{\mathcal{F}^{[2]}}(x^{[2]}),\mathcal{F}^{[2]})\to(\mathbf{Q}^{[1]},\mathcal{F}^{[2]});\ \mathbf{x}=(\mathbf{x}',\mathbf{x}'')\mapsto\mathbf{x}''$$

be the projections. Note that

$$(\overline{\mathcal{F}^{[1]}}(x^{[1]}),\mathcal{F}^{[2]})\simeq(X,T)\text{ and }(\mathbf{Q}^{[1]},\mathcal{F}^{[2]})\simeq(X\times X,\mathcal{G}^{[1]}).$$

Then $(p_2)_*(\lambda)$ is a $\mathcal{G}^{[1]}$-invariant measure of $\mathbf{Q}^{[1]}=X^{[1]}$. By subsection 4.1, $(p_2)_*(\lambda)=\mu^{[1]}=\mu\times\mu$. Hence let

$$\lambda=\int_{X^2}\lambda_{(x,y)}\times\delta_{(x,y)}\ d(\mu\times\mu)(x,y) \tag{4.16}$$

be the disintegration of λ over $\mu^{[1]}$. Since λ is $T_2^{[2]}=\mathrm{id}^{[1]}\times T^{[1]}$-invariant, we have

$$\begin{aligned}\lambda &= \mathrm{id}^{[1]}\times T^{[1]}\lambda=\int_{X^2}\lambda_{(x,y)}\times T^{[1]}\delta_{(x,y)}\ d\mu\times\mu(x,y)\\ &= \int_{X^2}\lambda_{(x,y)}\times\delta_{T^{[1]}(x,y)}\ d\mu\times\mu(x,y)\\ &= \int_{X^2}\lambda_{(T^{[1]})^{-1}(x,y)}\times\delta_{(x,y)}\ d\mu\times\mu(x,y).\end{aligned}$$

The uniqueness of disintegration implies that

$$\lambda_{(T^{[1]})^{-1}(x,y)}=\lambda_{(x,y)},\quad \mu^{[1]}=\mu\times\mu\ a.e. \tag{4.17}$$

Define

$$F:(\mathbf{Q}^{[1]}=X^{[1]},T^{[1]})\longrightarrow M(X):\ (x,y)\mapsto\lambda_{(x,y)}.$$

By (4.17), F is a $T^{[1]}$-invariant $M(X)$-value function. Hence F is $\mathcal{I}^{[1]}$-measurable, and hence $\lambda_{(x,y)}=\lambda_{\phi(x,y)}=\lambda_s$, $\mu^{[1]}$ a.e., where ϕ is defined in (4.10).

Thus by (4.16) one has that

$$\begin{aligned}\lambda &= \int_{X^2} \lambda_{(x,y)} \times \delta_{(x,y)}\ d\mu\times\mu(x,y) = \int_{X^2} \lambda_{\phi(x,y)} \times \delta_{(x,y)}\ d\mu\times\mu(x,y)\\ &= \int_{Z_1}\int_{X^2} \lambda_s \times \delta_{(x,y)}\ d\mu_s(x,y) d\mu_1(s)\\ &= \int_{Z_1} \lambda_s \times \Big(\int_{X^2} \delta_{(x,y)}\ d\mu_s(x,y)\Big) d\mu_1(s)\\ &= \int_{Z_1} \lambda_s \times \mu_s\ d\mu_1(s)\end{aligned}$$

Let $\pi_1^{[2]} : (\overline{\mathcal{F}^{[2]}}(x^{[2]}), \mathcal{F}^{[2]}) \longrightarrow (\overline{\mathcal{F}^{[2]}}((\pi_1(x))^{[2]}), \mathcal{F}^{[2]})$ be the natural factor map. By Theorem 2.13, $\overline{\mathcal{F}^{[2]}}((\pi_1(x))^{[2]})$ is uniquely ergodic. Hence

$$\pi_{1*}^{[2]}(\lambda) = \int_{Z_1} \mu_1 \times \mu_{1,s}\ d\mu_1(s) = \mu_1^3.$$

And

$$\pi_{1*}(\lambda_s) = \mu_1, \text{ and } (\pi_1\times\pi_1)_*(\mu_s) = \mu_{1,s}.$$

Note that we have that

$$(p_1)_*(\lambda) = \mu, \text{ and } (p_2)_*(\lambda) = \mu^{[1]} = \mu\times\mu,$$

and hence we have

$$\mu = \int_{Z_1} \lambda_s\ d\mu_1(s).$$

Let $\mu = \int_{Z_1} \nu_s\ d\mu_1(s)$ be the disintegration of μ over μ_1. Hence by the uniqueness of disintegration, we have that $\lambda_s = \nu_s$, μ_1 a.e.. Thus

$$\lambda = \int_{Z_1} \lambda_s \times \mu_s\ d\mu_1(s) = \int_{Z_1} \nu_s \times \mu_s\ d\mu_1(s).$$

That is, $(\overline{\mathcal{F}^{[2]}}(x^{[2]}), \mathcal{F}^{[2]})$ is uniquely ergodic. The proof is completed.

4.4. **General case.** In this section we prove Theorem A in the general case. We prove it by induction on d. $d = 1$ is showed in subsection 4.1. Now we assume d and show the case when $d+1$.

4.4.1. *Notations.* Recall that $\mathcal{I}^{[d]}$ is the $T^{[d]}$-invariant σ-algebra of $(X^{[d]}, \mu^{[d]}, T^{[d]})$ and

$$\mu^{[d+1]} = \mu^{[d]} \underset{\mathcal{I}^{[d]}}{\times} \mu^{[d]}.$$

Let

$$(X^{[d]}, \mu^{[d]}) \xrightarrow{\phi} (\Omega_d, \mathcal{I}^{[d]}, P_d);\ \mathbf{x} \longrightarrow \phi(\mathbf{x}) \tag{4.18}$$

be the factor map. Let

$$\mu^{[d]} = \int_{\Omega_d} \mu_\omega^{[d]}\ dP_d(\omega) \tag{4.19}$$

denote the ergodic decomposition of $\mu^{[d]}$ under $T^{[d]}$. Then by definition

$$\mu^{[d+1]} = \int_{\Omega_d} \mu^{[d]}_\omega \times \mu^{[d]}_\omega \, dP_d(\omega). \tag{4.20}$$

4.4.2. *A property about Z_d.*

Proposition 4.3. [9, Proposition 4.7.] *Let $d \geq 1$ be an integer.*

1. *As a joining of 2^d copies of (X, μ), $(X^{[d]}, \mu^{[d]})$ is relatively independent over the joining $(Z^{[d]}_{d-1}, \mu^{[d]}_{d-1})$ of 2^d copies of (Z_{d-1}, μ_{d-1}).*
2. *Z_d is the smallest factor Y of X so that the σ-algebra $\mathcal{I}^{[d]}$ is measurable with respect to $Y^{[d]}$.*

We say that a factor map $\pi : (X, \mathcal{X}, \mu, T) \to (Y, \mathcal{Y}, \nu, T)$ is an *ergodic* extension if every T-invariant $\mathcal{X}$-measurable function is $\mathcal{Y}$-measurable, i.e. $\mathcal{I}(X, T) \subset \mathcal{Y}$. Thus Proposition 4.3 implies that

$$\pi^{[d]}_d : (X^{[d]}, \mu^{[d]}, T^{[d]}) \to (Z^{[d]}_d, \mu^{[d]}_d, T^{[d]})$$

is $T^{[d]}$-ergodic. That means that $\mathcal{I}^{[d]}(X) = \mathcal{I}^{[d]}(Z_d)$, and hence $(\Omega_d(X), \mathcal{I}^{[d]}(X), P_d) = (\Omega_d(Z_d), \mathcal{I}^{[d]}(Z_d), P_d)$. So we can denote the ergodic decomposition of $\mu^{[d]}_d$ under $T^{[d]}$ by

$$\mu^{[d]}_d = \int_{\Omega_d} \mu^{[d]}_{d,\omega} \, dP_d(\omega). \tag{4.21}$$

Then by definition

$$\mu^{[d+1]}_d = \int_{\Omega_d} \mu^{[d]}_{d,\omega} \times \mu^{[d]}_{d,\omega} \, dP_d(\omega). \tag{4.22}$$

This property is crucial in the proof. Combining (4.18) and (4.21), one has factor maps

$$(X^{[d]}, \mu^{[d]}) \xrightarrow{\pi^{[d]}_d} (Z^{[d]}_d, \mu^{[d]}_d) \xrightarrow{\psi} (\Omega_d, P_d) \tag{4.23}$$

Note that $\phi = \psi \circ \pi^{[d]}_d$.

4.4.3. *$\mathcal{G}$-action.* Now we assume that Theorem A(2) holds for $d \geq 1$. In this subsection we show the existence of $\mathcal{G}^{[d+1]}$-model.

Let $\pi_d : X \to Z_d$ be the factor map from X to its d-step nilfactor Z_d. By definition, Z_d may be regarded as a topological system in the natural way. By Weiss's Theorem, there is a uniquely ergodic model $(\hat{X}, \hat{\mathcal{X}}, \hat{\mu}, T)$ for $(X, \mathcal{X}, \mu, T)$ and a factor map $\hat{\pi}_d : \hat{X} \to Z_d$ which is a model for $\pi_d : X \to Z_d$.

$$\begin{array}{ccc} X & \longrightarrow & \hat{X} \\ {\scriptstyle \pi_d}\downarrow & & \downarrow{\scriptstyle \hat{\pi}_d} \\ Z_d & \longrightarrow & Z_d \end{array}$$

Hence for simplicity, we may assume that $(\hat{X}, \hat{\mathcal{X}}, \hat{\mu}, T) = (X, \mathcal{X}, \mu, T)$ and $\pi_d = \hat{\pi}_d$. Now we show that $(\mathbf{Q}^{[d+1]}(X), \mu^{[d+1]}, \mathcal{G}^{[d+1]})$ is uniquely ergodic.

Let λ be a $\mathcal{G}^{[d+1]}$-invariant measure of $\mathbf{Q}^{[d+1]} = \mathbf{Q}^{[d+1]}(X)$. Let

$$p_1 : (\mathbf{Q}^{[d+1]}, \mathcal{G}^{[d+1]}) \to (\mathbf{Q}^{[d]}, \mathcal{G}^{[d+1]});\ \mathbf{x} = (\mathbf{x}', \mathbf{x}'') \mapsto \mathbf{x}'$$
$$p_2 : (\mathbf{Q}^{[d+1]}, \mathcal{G}^{[d+1]}) \to (\mathbf{Q}^{[d]}, \mathcal{G}^{[d+1]});\ \mathbf{x} = (\mathbf{x}', \mathbf{x}'') \mapsto \mathbf{x}''$$

be the projections. Then $(p_2)_*(\lambda)$ is a $\mathcal{G}^{[d+1]}$-invariant measure of $\mathbf{Q}^{[d]}$. Note that $\mathcal{G}^{[d+1]}$ acts on $\mathbf{Q}^{[d]}$ as $\mathcal{G}^{[d]}$ actions. By the induction hypothesis, $(p_2)_*(\lambda) = \mu^{[d]}$. Hence let

$$\lambda = \int_{\mathbf{Q}^{[d]}} \lambda_{\mathbf{x}} \times \delta_{\mathbf{x}}\ d\mu^{[d]}(\mathbf{x}) \tag{4.24}$$

be the disintegration of λ over $\mu^{[d]}$. Since λ is $T_{d+1}^{[d+1]} = \mathrm{id}^{[d]} \times T^{[d]}$-invariant, we have

$$\begin{aligned} \lambda &= \mathrm{id}^{[d]} \times T^{[d]} \lambda = \int_{\mathbf{Q}^{[d]}} \lambda_{\mathbf{x}} \times T^{[d]} \delta_{\mathbf{x}}\ d\mu^{[d]}(\mathbf{x}) \\ &= \int_{\mathbf{Q}^{[d]}} \lambda_{\mathbf{x}} \times \delta_{T^{[d]}(\mathbf{x})}\ d\mu^{[d]}(\mathbf{x}) \\ &= \int_{\mathbf{Q}^{[d]}} \lambda_{(T^{[d]})^{-1}(\mathbf{x})} \times \delta_{\mathbf{x}}\ d\mu^{[d]}(\mathbf{x}). \end{aligned}$$

The uniqueness of disintegration implies that

$$\lambda_{(T^{[d]})^{-1}(\mathbf{x})} = \lambda_{\mathbf{x}}, \quad \mu^{[d]}\ a.e.\ \mathbf{x} \in \mathbf{Q}^{[d]}. \tag{4.25}$$

Define

$$F : (\mathbf{Q}^{[d]}, T^{[d]}) \longrightarrow M(X^{[d]}) :\ \mathbf{x} \mapsto \lambda_{\mathbf{x}}.$$

By (4.25), F is a $T^{[d]}$-invariant $M(X^{[d]})$-value function. Hence F is $\mathcal{I}^{[d]}$-measurable, and hence $\lambda_{\mathbf{x}} = \lambda_{\phi(\mathbf{x})},\ \mu^{[d]}$ a.e., where ϕ is defined in (4.18).

Thus by (4.24) one has that

$$\begin{aligned} \lambda &= \int_{\mathbf{Q}^{[d]}} \lambda_{\mathbf{x}} \times \delta_{\mathbf{x}}\ d\mu^{[d]}(\mathbf{x}) = \int_{\mathbf{Q}^{[d]}} \lambda_{\phi(\mathbf{x})} \times \delta_{\mathbf{x}}\ d\mu^{[d]}(\mathbf{x}) \\ &= \int_{\Omega_d} \int_{\mathbf{Q}^{[d]}} \lambda_\omega \times \delta_{\mathbf{x}}\ d\mu_\omega^{[d]}(\mathbf{x}) dP_d(\omega) \\ &= \int_{\Omega_d} \lambda_\omega \times \Big(\int_{\mathbf{Q}^{[d]}} \delta_{\mathbf{x}}\ d\mu_\omega^{[d]}(\mathbf{x}) \Big) dP_d(\omega) \\ &= \int_{\Omega_d} \lambda_\omega \times \mu_\omega^{[d]}\ dP_d(\omega) \end{aligned}$$

Let $\pi_d^{[d+1]} : (\mathbf{Q}^{[d+1]}(X), \mathcal{G}^{[d+1]}) \longrightarrow (\mathbf{Q}^{[d+1]}(Z_d), \mathcal{G}^{[d+1]})$ be the natural factor map. By Theorem 2.13, $(\mathbf{Q}^{[d+1]}(Z_d), \mu_d^{[d+1]})$ is uniquely ergodic. Hence

$$\pi_{d*}^{[d+1]}(\lambda) = \mu_d^{[d+1]} = \int_{\Omega_d} \mu_{d,\omega}^{[d]} \times \mu_{d,\omega}^{[d]}\ dP_d(\omega).$$

So

$$\pi_{d*}^{[d]}(\lambda_\omega) = \pi_{d*}^{[d]}(\mu_\omega^{[d]}) = \mu_{d,\omega}^{[d]}. \tag{4.26}$$

Note that we have that

$$(p_1)_*(\lambda) = (p_2)_*(\lambda) = \mu^{[d]},$$

and hence we have

$$\mu^{[d]} = \int_{\Omega_d} \lambda_\omega \ dP_d(\omega) = \int_{\Omega_d} \mu_\omega^{[d]} \ dP_d(\omega). \tag{4.27}$$

But by (4.26) and (4.23) we have

$$\phi_*(\lambda_\omega) = \phi_*(\mu_\omega^{[d]}) = \psi_*(\mu_{d,\omega}^{[d]}) = \delta_\omega.$$

Hence by the uniqueness of disintegration and (4.27), we have that $\lambda_\omega = \mu_\omega^{[d]}$, P_d a.e. $\omega \in \Omega_d$. Thus we have

$$\lambda_{\mathbf{Q};d+1} = \int_{\Omega_d} \lambda_\omega \times \mu_\omega^{[d]} \ dP_d(\omega) = \int_{\Omega_d} \mu_\omega^{[d]} \times \mu_\omega^{[d]} \ dP_d(\omega) = \mu^{[d+1]}. \tag{4.28}$$

That is, $(\mathbf{Q}^{[d+1]}, \mu^{[d+1]}, \mathcal{G}^{[d+1]})$ is uniquely ergodic. The proof of Theorem A(2) for $\mathcal{G}$ is completed.

4.4.4. *$\mathcal{F}$-actions.* Now we assume that Theorem A(1) holds for $d \geq 1$. In this subsection we show the existence of $\mathcal{F}^{[d+1]}$-model. We use the same model as in the previous subsection.

Let λ be a $\mathcal{F}^{[d+1]}$-invariant measure of $\overline{\mathcal{F}^{[d+1]}}(x^{[d+1]})$. Let

$$p_1 : (\overline{\mathcal{F}^{[d+1]}}(x^{[d+1]}), \mathcal{F}^{[d+1]}) \to (\overline{\mathcal{F}^{[d]}}(x^{[d]}), \mathcal{F}^{[d+1]});\ \mathbf{x} = (\mathbf{x}', \mathbf{x}'') \mapsto \mathbf{x}'$$
$$p_2 : (\overline{\mathcal{F}^{[d+1]}}(x^{[d+1]}), \mathcal{F}^{[d+1]}) \to (\mathbf{Q}^{[d]}, \mathcal{F}^{[d+1]});\ \mathbf{x} = (\mathbf{x}', \mathbf{x}'') \mapsto \mathbf{x}''$$

be the projections. Note that

$$(\overline{\mathcal{F}^{[d]}}(x^{[d]}), \mathcal{F}^{[d+1]}) \simeq (\overline{\mathcal{F}^{[d]}}(x^{[d]}), \mathcal{F}^{[d]}) \text{ and } (\mathbf{Q}^{[d]}, \mathcal{F}^{[d+1]}) \simeq (\mathbf{Q}^{[d]}, \mathcal{G}^{[d]}).$$

Then $(p_2)_*(\lambda)$ is a $\mathcal{G}^{[d]}$-invariant measure of $\mathbf{Q}^{[d]}$. By subsection 4.4.3, $(p_2)_*(\lambda) = \mu^{[d]}$. Hence let

$$\lambda = \int_{\mathbf{Q}^{[d]}} \lambda_{\mathbf{x}} \times \delta_{\mathbf{x}} \ d\mu^{[d]}(\mathbf{x}) \tag{4.29}$$

be the disintegration of λ over $\mu^{[d]}$. Since λ is $T_{d+1}^{[d+1]} = \mathrm{id}^{[d]} \times T^{[d]}$-invariant, we have

$$\begin{aligned} \lambda &= \mathrm{id}^{[d]} \times T^{[d]}\lambda = \int_{\mathbf{Q}^{[d]}} \lambda_{\mathbf{x}} \times T^{[d]}\delta_{\mathbf{x}} \ d\mu^{[d]}(\mathbf{x}) \\ &= \int_{\mathbf{Q}^{[d]}} \lambda_{\mathbf{x}} \times \delta_{T^{[d]}(\mathbf{x})} \ d\mu^{[d]}(\mathbf{x}) \\ &= \int_{\mathbf{Q}^{[d]}} \lambda_{(T^{[d]})^{-1}(\mathbf{x})} \times \delta_{\mathbf{x}} \ d\mu^{[d]}(\mathbf{x}). \end{aligned}$$

The uniqueness of disintegration implies that

$$\lambda_{(T^{[d]})^{-1}(\mathbf{x})} = \lambda_{\mathbf{x}}, \quad \mu^{[d]} \ a.e. \tag{4.30}$$

Define

$$F : \mathbf{Q}^{[d]} \longrightarrow M(\overline{\mathcal{F}^d}(x^{[d]})) : \ \mathbf{x} \mapsto \lambda_{\mathbf{x}}.$$

By (4.30), F is a $T^{[d]}$-invariant $M(\overline{\mathcal{F}^d}(x^{[d]}))$-value function. Hence F is $\mathcal{I}^{[d]}$-measurable, and hence $\lambda_{\mathbf{x}} = \lambda_{\phi(\mathbf{x})}$, $\mu^{[d]}$ a.e. $\mathbf{x} \in \mathbf{Q}^{[d]}$, where ϕ is defined in (4.18).

Thus by (4.29) one has that

$$\begin{aligned}
\lambda &= \int_{\mathbf{Q}^{[d]}} \lambda_{\mathbf{x}} \times \delta_{\mathbf{x}} \ d\mu^{[d]}(\mathbf{x}) = \int_{\mathbf{Q}^{[d]}} \lambda_{\phi(\mathbf{x})} \times \delta_{\mathbf{x}} \ d\mu^{[d]}(\mathbf{x}) \\
&= \int_{\Omega_d} \int_{\mathbf{Q}^{[d]}} \lambda_\omega \times \delta_{\mathbf{x}} \ d\mu^{[d]}_\omega(\mathbf{x}) dP_d(\omega) \\
&= \int_{\Omega_d} \lambda_\omega \times \Big(\int_{\mathbf{Q}^{[d]}} \delta_{\mathbf{x}} \ d\mu^{[d]}_\omega(\mathbf{x}) \Big) dP_d(\omega) \\
&= \int_{\Omega_d} \lambda_\omega \times \mu^{[d]}_\omega \ dP_d(\omega)
\end{aligned}$$

Since $(\overline{\mathcal{F}^d}(x^{[d]}), \mathcal{F}^{[d]})$ is uniquely ergodic by assumption, and we let $\nu_x^{[d]}$ be the unique measure. Then

$$(p_1)_*(\lambda) = \nu_x^{[d]}, \text{ and } (p_2)_*(\lambda) = \mu^{[d]},$$

and hence we have

$$\nu_x^{[d]} = \int_{\Omega_d} \lambda_\omega \ dP_d(\omega). \tag{4.31}$$

Note that we have a factor map $\pi_d^{[d]} : (\overline{\mathcal{F}^{[d]}}(x^{[d]}), \mathcal{F}^{[d]}, \nu_x^{[d]}) \to (\overline{\mathcal{F}^{[d]}}(\hat{x}^{[d]}), \mathcal{F}^{[d]}, \rho_{d,\hat{x}})$, where $\hat{x} = \pi_d(x)$ and $\rho_{d,\hat{x}}$ as in Theorem 2.13. For each $z \in \overline{\mathcal{F}^{[d]}}(\hat{x}^{[d]})$, let η_z be the unique $T^{[d]}$-invariant measure on $\overline{\mathcal{O}(z, T^{[d]})}$. Then the map

$$\overline{\mathcal{F}^{[d]}}(\hat{x}^{[d]}) \to M(\mathbf{Q}^{[d]}(Z_d)); \ z \mapsto \eta_z$$

is a measurable map. This fact follows from that $z \mapsto \frac{1}{N} \sum_{n<N} \delta_{T^n z}$ is continuous and $\frac{1}{N} \sum_{n<N} \delta_{T^n z}$ converges to η_z weakly. Hence we have

$$\mu_d^{[d]} = \int_{\overline{\mathcal{F}^{[d]}}(\hat{x}^{[d]})} \eta_z \ d\rho_{d,\hat{x}}(z). \tag{4.32}$$

In fact, it is easy to check that $\int_{\overline{\mathcal{F}^{[d]}}(\hat{x}^{[d]})} \eta_z \ d\rho_{d,\hat{x}}(z)$ is $\mathcal{G}^{[d]}$-invariant and hence it is equal to $\mu_d^{[d]}$ by the uniqueness. Note that (4.32) is the "ergodic decomposition" of $\mu_d^{[d]}$ under $T^{[d]}$, except that it happens that $\eta_z = \eta_{z'}$ for some $z \neq z'$. Hence via map ψ, we have a factor map

$$\Psi : (\overline{\mathcal{F}^{[d]}}(\hat{x}^{[d]}), \rho_{d,\hat{x}}) \to (\Omega_d, P_d).$$

And (4.32) can be rewritten as

$$\mu_d^{[d]} = \int_{\overline{\mathcal{F}^{[d]}}(\hat{x}^{[d]})} \eta_z \ d\rho_{d,\hat{x}}(z) = \int_{\Omega_d} \eta_\omega \ dP_d(\omega) = \int_{\Omega_d} \mu^{[d]}_{d,\omega} \ dP_d(\omega). \tag{4.33}$$

Since we have

$$(\overline{\mathcal{F}^{[d]}}(x^{[d]}), \nu_x^{[d]}) \xrightarrow{\pi_d^{[d]}} (\overline{\mathcal{F}^{[d]}}(\hat{x}^{[d]}), \rho_{d,\hat{x}}) \xrightarrow{\Psi} (\Omega_d, P_d) \tag{4.34}$$

we assume that

$$\nu_x^{[d]} = \int_{\Omega_d} \nu_\omega \, dP_d(\omega) \tag{4.35}$$

is the disintegration of $\nu_x^{[d]}$ over Ω_d.

Let $\pi_d^{[d+1]} : (\overline{\mathcal{F}^{[d+1]}}(x^{[d+1]}), \mathcal{F}^{[d+1]}) \longrightarrow (\overline{\mathcal{F}^{[d+1]}}(\hat{x}^{[d+1]}), \mathcal{F}^{[d+1]})$ be the natural factor map. By Theorem 2.13, $(\overline{\mathcal{F}^{[d+1]}}((\hat{x})^{[d+1]}), \rho_{d+1,\hat{x}})$ is uniquely ergodic. Let

$$(\pi_d^{[d+1]})_*(\lambda) = \rho_{d+1,\hat{x}} = \int_{\overline{\mathcal{F}^{[d]}}(\hat{x}^{[d]})} \delta_z \times \eta_z \, d\rho_{d,x}(z).$$

be the disintegration of $\rho_{d+1,x}$ over $\overline{\mathcal{F}^{[d]}}(\hat{x}^{[d]})$. By (4.34), we have

$$(\pi_d^{[d+1]})_*(\lambda) = \rho_{d+1,\hat{x}} = \int_{\overline{\mathcal{F}^{[d]}}(\hat{x}^{[d]})} \delta_z \times \eta_z \, d\rho_{d,x}(z) = \int_{\Omega_d} \rho_\omega \times \mu_{d,\omega}^{[d]} \, dP_d(\omega),$$

where $\rho_{d,\hat{x}} = \int_{\Omega_d} \rho_\omega \, dP_d(\omega)$ is the disintegration of $\rho_{d,\hat{x}}$ over P_d. Then

$$(\pi_d^{[d]})_*(\lambda_\omega) = \rho_\omega, \text{ and } (\pi_d^{[d]})_*(\mu_\omega^{[d]}) = \mu_{d,\omega}^{[d]}. \tag{4.36}$$

Since $(\pi_d^{[d]})_*(\nu_x^{[d]}) = \rho_{d,\hat{x}}$, by (4.35) we have $(\pi_d^{[d]})_*(\nu_\omega) = \rho_\omega$. Hence by the uniqueness of disintegration, we have that $\lambda_\omega = \nu_\omega$, P_d a.e.. Thus

$$\lambda_{\mathcal{F};d+1} = \lambda = \int_{\Omega_d} \lambda_\omega \times \mu_\omega^{[d]} \, dP_d(\omega) = \int_{\Omega_d} \nu_\omega \times \mu_\omega^{[d]} \, dP_d(\omega). \tag{4.37}$$

That is, λ is unique and hence $(\overline{\mathcal{F}^{[d+1]}}(x^{[d+1]}), \mathcal{F}^{[d+1]})$ is uniquely ergodic. The proof is completed. □

References

[1] I. Assani, *Pointwise convergence of ergodic averages along cubes*, J. Analyse Math. **110** (2010), 241–269.

[2] V. Bergelson. *The multifarious Poincare recurrence theorem*, Descriptive set theory and dynamical systems. London Math. Soc. Lecture Note Series 277, Cambridge Univ. Press, Cambridge, (2000), 31–57.

[3] V. Bergelson, *Combinatorial and Diophantine applications of ergodic theory*, Appendix A by A. Leibman and Appendix B by Anthony Quas and Máté Wierdl. Handbook of dynamical systems. Vol. 1B, 745–869, Elsevier B. V., Amsterdam, 2006.

[4] Q. Chu and N. Frantzikinakis, *Pointwise convergence for cubic and polynomial ergodic averages of non-commuting transformations*, Ergod. Th. and Dynam. Sys. **32** (2012), 877–897.

[5] H. Furstenberg, *Recurrence in ergodic theory and combinatorial number theory*, M. B. Porter Lectures. Princeton University Press, Princeton, N.J., 1981.

[6] H. Furstenberg and B. Weiss, *On almost* 1-1 *extensions*, Israel J. Math. 65 (1989), no. 3, 311–322.

[7] E. Glasner, *Ergodic theory via joinings*, Mathematical Surveys and Monographs, 101. American Mathematical Society, Providence, RI, 2003. xii+384 pp.

[8] E. Glasner and B. Weiss, *Strictly ergodic, uniform positive entropy models*, Bull. Soc. Math. France 122 (1994), no. 3, 399–412.

[9] B. Host and B. Kra, *Nonconventional averages and nilmanifolds*, Ann. of Math., 161 (2005) 398–488.

[10] B. Host and B. Kra, *Uniformity norms on l^∞ and applications*, J. Anal. Math., **108** (2009), 219–276.

[11] B. Host, B. Kra and A. Maass, *Nilsequences and a structure theory for topological dynamical systems*, Advances in Mathematics, 224 (2010) 103–129.

[12] W. Huang, S. Shao and X. Ye, *Nil Bohr-sets and almost automorphy of higher order*, to be published in *Memoirs of the AMS*.

[13] W. Huang, S. Shao and X. Ye, *Pointwise convergence of multiple ergodic averages and strictly ergodic models*, preprint. arXiv:1406.5930

[14] R.I. Jewett, *The prevalence of uniquely ergodic systems*, J. Math. Mech. 19 1969/1970 717–729.

[15] W. Krieger, *On unique ergodicity*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory, pp. 327–346, Univ. California Press, Berkeley, Calif., 1972.

[16] E. Lehrer, *Topological mixing and uniquely ergodic systems*, Israel J. Math. 57 (1987), no. 2, 239–255.

[17] E. Lindenstrauss, *Pointwise theorems for amenable groups.* Invent. Math. **146** (2001), 259–295.

[18] S. Shao and X. Ye, *Regionally proximal relation of order d is an equivalence one for minimal systems and a combinatorial consequence*, Adv. in Math., **231**(2012), 1786–1817.

[19] B. Weiss, *Strictly ergodic models for dynamical systems*, Bull. Amer. Math. Soc. (N.S.) **13** (1985), 143–146.

[20] B. Weiss, *Countable generators in dynamics – Universal minimal models*, Contemporary Mathematics 94, (1989), 321–326.

[21] B. Weiss, *Single orbit dynamics*, CBMS Regional Conference Series in Mathematics, 95. American Mathematical Society, Providence, RI, 2000. x+113 pp.

[22] T. Ziegler, *Universal characteristic factors and Furstenberg averages.* J. Amer. Math. Soc. 20 (2007), no. 1, 53–97.

Department of Mathematics, University of Science and Technology of China, Hefei, Anhui, 230026, P.R. China.

E-mail address: wenh@mail.ustc.edu.cn

E-mail address: songshao@ustc.edu.cn

E-mail address: yexd@ustc.edu.cn