# Multiple recurrence for two commuting transformations 

Qing Chu

## To cite this version:

Qing Chu. Multiple recurrence for two commuting transformations. 2009. hal-00441767

## HAL Id: hal-00441767 <br> https://hal.science/hal-00441767

Preprint submitted on 17 Dec 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# MULTIPLE RECURRENCE FOR TWO COMMUTING TRANSFORMATIONS 

QING CHU


#### Abstract

This paper is devoted to a study of the multiple re－ currence of two commuting transformations．We derive a result which is similar but not identical to that of one single transfor－ mation established by Bergelson，Host and Kra．We will use the machinery of＂magic systems＂established recently by B．Host for the proof．


## 1．Introduction

1．1．History and results．Let $(X, \mathcal{X}, \mu, T)$ be an invertible measure preserving system，and $A$ be a set of positive measure．The Khint－ chine＇s Recurrence Theorem［14］states that for every $\epsilon>0$ ，the set

$$
\left\{n \in \mathbb{Z}: \mu\left(A \cap T^{n} A\right)>\mu(A)^{2}-\epsilon\right\}
$$

is syndetic．More recently，Furstenberg［9］proved a Multiple Recur－ rence Theorem，showing that under the same assumptions，the set

$$
\left\{n \in \mathbb{Z}: \mu\left(A \cap T^{n} A \cap T^{2 n} A \cap \cdots \cap T^{k n} A\right)>0\right\}
$$

is syndetic for every integer $k \geq 1$ ．Aiming at a simultaneous exten－ sion of Khintchine＇s and Furstenberg＇s Recurrence theorems，Bergelson， Host and Kra［⿴囗十⺝刂 established the following result：

Theorem（Bergelson，Host，Kra）．Let $(X, \mathcal{X}, \mu, T)$ be an ergodic in－ vertible measure preserving system and $A \in \mathcal{X}$ with $\mu(A)>0$ ．Then for every $\epsilon>0$ ，the sets

$$
\left\{n \in \mathbb{Z}: \mu\left(A \cap T^{n} A \cap T^{2 n} A\right)>\mu(A)^{3}-\epsilon\right\}
$$

and

$$
\left\{n \in \mathbb{Z}: \mu\left(A \cap T^{n} A \cap T^{2 n} A \cap T^{3 n} A\right)>\mu(A)^{4}-\epsilon\right\}
$$

are syndetic．

[^0]We recall that a subset $E$ of $\mathbb{Z}$ is said to be syndetic if there exists an integer $N>0$ such that $E \cap[M, M+N) \neq \emptyset$ for every $M \in \mathbb{Z}$.

It was shown in [1] that an analogous result fails both if we remove the assumption of ergodicity and if for longer arithmetic progressions.

Furstenberg and Katznelson [11] generalized Furstenberg's Recurrence theorem to commuting transformations. It is therefore natural to ask the question if a result analogous to the above theorem can be established for commuting transformations. We prove the following result regarding the case of two transformations:
Theorem 1.1. Let $(X, \mathcal{X}, \mu)$ be a probability space, and $T_{1}, T_{2}$ be two commuting invertible measure preserving transformations. Assume that $\left(X, \mathcal{X}, \mu, T_{1}, T_{2}\right)$ is ergodic. Let $A \in \mathcal{X}$ with $\mu(A)>0$. Then for every $\epsilon>0$, the set

$$
\left\{n \in \mathbb{Z}: \mu\left(A \cap T_{1}^{n} A \cap T_{2}^{n} A\right)>\mu(A)^{4}-\epsilon\right\}
$$

is syndetic.
We remark that, by the same counterexample as in [4], the hypothesis of ergodicity is necessary for the theorem.

The fundamental difference with the case of a single transformation is that the exponent 4 can not be replaced by 3 :

Theorem 1.2. For every $0<c \leq 1$, there exist a probability space $(X, \mathcal{X}, \mu)$, with two commuting invertible measure preserving transformations $T_{1}, T_{2}$ such that $\left(X, \mathcal{X}, \mu, T_{1}, T_{2}\right)$ is ergodic, and a measurable set $A \in \mathcal{X}$, with $\mu(A)>0$, such that

$$
\mu\left(A \cap T_{1}^{n} A \cap T_{2}^{n} A\right)<c \mu(A)^{3}
$$

for every integer $n \neq 0$.
However, we have the exponent 3 for some class of systems:
Theorem 1.3. Let $\left(Y_{1}, \mathcal{Y}_{1}, \nu_{1}, S_{1}\right)$ and $\left(Y_{2}, \mathcal{Y}_{2}, \nu_{2}, S_{2}\right)$ be two ergodic invertible measure preserving systems. Let $(X, \mathcal{X}, \mu)$ be the product measure space $\left(Y_{1} \times Y_{2}, \mathcal{Y}_{1} \otimes \mathcal{Y}_{2}, \nu_{1} \times \nu_{2}\right)$, and let $T_{1}=S_{1} \times \mathrm{Id}$, $T_{2}=$ Id $\times S_{2}$. Let $A \in \mathcal{X}$ with $\mu(A)>0$. Then for every $\epsilon>0$, the set

$$
\left\{n \in \mathbb{Z}: \mu\left(A \cap T_{1}^{n} A \cap T_{2}^{n} A\right)>\mu(A)^{3}-\epsilon\right\}
$$

is syndetic.
We recall a definition:
Definition 1.4. The upper Banach density of a subset $E$ of $\mathbb{Z}^{2}$ is:

$$
d^{*}(E)=\limsup _{\substack{N_{1}-M_{1} \rightarrow \infty \\ N_{2}-M_{2} \rightarrow \infty}} \frac{\left|E \cap\left[M_{1}, N_{1}\right) \times\left[M_{2}, N_{2}\right)\right|}{\left(N_{1}-M_{1}\right) \times\left(N_{2}-M_{2}\right)} .
$$

Using a variation（see（⿴囗⿱一一⿰亻⿴ ）of Multidimensional Furstenberg＇s Corre－ spondence Principle［10］and Theorem［1．1，we deduce：

Corollary 1．5．Let $E \subset \mathbb{Z}^{2}$ be a subset with positive upper Banach density．Then for every $\epsilon>0$ ，the set

$$
\left\{n \in \mathbb{Z}: \quad d^{*}(E \cap(E+(n, 0)) \cap(E+(0, n)))>d^{*}(E)^{4}-\epsilon\right\}
$$

is syndetic．
1．2．Questions．We address here some questions which are related to this paper and remain open．

Question 1：A natural question is how about the case of three commuting transformations．We ask if there exists some integer $s$ ，such that，for every ergodic system $\left(X, \mathcal{X}, \mu, T_{1}, T_{2}, T_{3}\right)$ and every set $A \in \mathcal{X}$ with $\mu(A)>0$ ，the set $\left\{n \in \mathbb{Z}: \mu\left(A \cap T_{1}^{n} A \cap T_{2}^{n} A \cap T_{3}^{n} A\right)>\mu(A)^{s}-\epsilon\right\}$ is syndetic．

Question 2：The Polynomial Recurrence Theorem was proved by Bergelson and Leibman［3．Frantzikinakis and Kra［8］provided a more precise result for any family of linearly independent integer polynomi－ als：

Theorem（Frantzikinakis，Kra）．Let $(X, \mathcal{X}, \mu, T)$ be an invertible mea－ sure preserving system，and $p_{1}, \ldots, p_{k}$ be linearly independent integer polynomials with $p_{i}(0)=0$ for $i=1, \ldots, k$ ，and $A \in \mathcal{X}$ ．Then for every $\epsilon>0$ ，the set

$$
\left\{n \in \mathbb{Z}: \quad \mu\left(A \cap T^{p_{1}(n)} A \cap \cdots \cap T^{p_{k}(n)} A\right) \geq \mu(A)^{k+1}-\epsilon\right\}
$$

is syndetic．
Can this result be generalized for commuting transformations？Let （ $X, \mathcal{X}, \mu$ ）be a probability space，and $T_{1}, \ldots, T_{k}$ be commuting invert－ ible measure preserving transformations．Let $p_{1}, \ldots, p_{k}$ be as in the above theorem，and $A \in \mathcal{X}$ ．It is true that the set $\{n \in \mathbb{Z}: \mu(A \cap$ $\left.\left.T_{1}^{p_{1}(n)} A \cap \cdots \cap T_{k}^{p_{k}(n)} A\right) \geq \mu(A)^{k+1}-\epsilon\right\}$ is syndetic？Very recently，Chu， Frantzikinakis，and Host［6］gave an affirmative answer to this question when $p_{i}(n)=n^{d_{i}}, i=1, \ldots, k$ for distinct positive integers $d_{1}, \ldots, d_{k}$ ．

## 1．3．Conventions and notation．

1．3．1．As usual，we can restrict to the case that all the probability spaces that we deal with are standard．

In general，we write $(X, \mu)$ for a probability space，omitting the $\sigma$－ algebra．When needed，the $\sigma$－algebra of a the probability space $(X, \mu)$ is written $\mathcal{X}$ ．

We implicitly assume that the term "bounded function" means realvalued, bounded and measurable.

If $S$ is a measure-preserving transformation of a probability space $(X, \mathcal{X}, \mu)$ then we write $\mathcal{I}(S)$ for the sub- $\sigma$-algebra of $\mathcal{X}$ consisting in $S$-invariant sets.

Let $\left(X_{1}, \mathcal{X}_{1}, \mu_{1}\right),\left(X_{2}, \mathcal{X}_{2}, \mu_{2}\right)$ be two probability spaces. Let $f_{1} \in$ $L^{\infty}\left(\mu_{1}\right)$, and $f_{2} \in L^{\infty}\left(\mu_{2}\right)$, we denote by $f_{1} \otimes f_{2}$ the function on $X_{1} \times X_{2}$ given by $f_{1} \otimes f_{2}\left(x_{1}, x_{2}\right)=f_{1}\left(x_{1}\right) f_{2}\left(x_{2}\right)$.
1.3.2. Throughout this paper, by a system, we mean a probability space $(X, \mathcal{X}, \mu)$ endowed with a single or several commuting measure preserving invertible transformations.

For a system $\left(X, \mu, T_{1}, T_{2}\right)$, a factor is a system $\left(Y, \nu, S_{1}, S_{2}\right)$ and a measurable map $\pi: X \rightarrow Y$ such that the image $\pi(\mu)$ of $\mu$ under $\pi$ is equal to $\nu$ and $S_{i} \circ \pi=\pi \circ T_{i}, i=1,2, \mu$-a.e.

If $f$ is an integrable function on $X$, we write $\mathbb{E}(f \mid Y)$ for the function on $Y$ defined by

$$
\text { for all } g \in L^{\infty}(\nu), \quad \int_{X} f \cdot g \circ \pi d \mu=\int_{Y} \mathbb{E}(f \mid Y) \cdot g d \nu .
$$

1.3.3. We say that the averages of some sequence $\left(a_{n}\right)$ converge to some limit $L$, and we write:

$$
\lim _{N-M \rightarrow \infty} \frac{1}{N-M} \sum_{n \in[M, N)} a_{n}=L
$$

if the averages of $a_{n}$ on any sequences of intervals [ $M_{i}, N_{i}$ ) whose lengths $N_{i}-M_{i}$ tend to infinity converge to $L$ :

$$
\lim _{i \rightarrow \infty} \frac{1}{N_{i}-M_{i}} \sum_{n \in\left[M_{i}, N_{i}\right)} a_{n}=L
$$

For $t \in \mathbb{T}$, we use the standard notation $e(n t)=\exp (2 \pi i n t)$.

### 1.4. Methods.

1.4.1. The first ingredient used in the proofs of Theorem 1.1 and Theorem 1.3 is a method introduced by Frantzikinakis [7]: in order to show some sequence $I_{n}$ is large for $n$ in a syndetic set, we show that the average of $I_{n}$ over an appropriately chosen sequence of intervals converges to a large limit. More precisely, we introduce an ergodic rotation $(Z, \alpha)$, and we show that for some well chosen non-negative continuous function $\chi$ on $Z$, the weighted average of $\chi(n \alpha) I_{n}$ converges to a large limit. This strategy is sufficient for the proof of Theorem 1.3 (Section 2 ).

However, for the general case (Theorem 1.1), before using the first ingredient, we need first to make some reduction. We will use more elaborated tools such as the machinery of "magic systems" introduced recently by Host [12].

Ever since Tao [17] proved the norm convergence of multiple ergodic averages with several commuting transformations of the form

$$
\begin{equation*}
\frac{1}{N} \sum_{n=1}^{N} T_{1}^{n} f_{1} \cdot \ldots \cdot T_{d}^{n} f_{d} \tag{1}
\end{equation*}
$$

several other proofs were given with different approaches by Austin [1], Host [12] and Towsner [18]. Among these proofs, those by Austin and Host were proceeded by building an extension of the original system with good properties, called "pleasant" system (using terminology from (1]) and "magic" system (using terminology from [12]). We will follow this idea of building a suitable extension system.

We first prove that every ergodic system has an ergodic magic extension. Then we use this result three times, and we get an ergodic magic extension. We consider our problem on this extension system. By using the properties of magic systems related to the convergence of multiple averages, we are reduced to consider our problem on a factor of this extension system. At last, we give a description of this factor, and we find that we are in a situation very similar to that of the product case.

We will review "magic systems" and the corresponding properties needed in Section 3, and give the proof of Theorem 1.1 in Section 4.
1.4.2. Another important ingredient is the following inequality. This inequality was proved in a particular case by Atkinson, Watterson, and Moran [2], and it is related to a class of inequalities studied by Sidorenko ( [15], [16]).

Lemma 1.6. Let $(X, \mathcal{X}, \mu)$ be a probability space, $k \geq 1$ be an integer, and $\mathcal{X}_{1}, \mathcal{X}_{2}, \ldots, \mathcal{X}_{k}$ be $k$ sub- $\sigma$-algebras of $\mathcal{X}$. For any bounded nonnegative function $f$ on $X$, we have

$$
\begin{equation*}
\int f \cdot \prod_{i=1}^{k} \mathbb{E}\left(f \mid \mathcal{X}_{i}\right) d \mu \geq\left(\int f d \mu\right)^{k+1} \tag{2}
\end{equation*}
$$

Proof. We can restrict to the case that the function $f$ is bounded below by some positive constant $\epsilon$. Indeed, for the general case, it suffices to apply the inequality to the function $f+\epsilon$ and to take the limit of both sides when $\epsilon$ tends to 0 .

We write

$$
\begin{equation*}
f=\left(f \cdot \prod_{i=1}^{k} \mathbb{E}\left(f \mid \mathcal{X}_{i}\right)\right)^{1 / k+1} \cdot \prod_{i=1}^{k}\left(\frac{f}{\mathbb{E}\left(f \mid \mathcal{X}_{i}\right)}\right)^{1 / k+1} \tag{3}
\end{equation*}
$$

Let $J$ be the integral on the left hand side of (21). By the Hölder Inequality and (3),

$$
\begin{equation*}
\left(\int f d \mu\right)^{k+1} \leq J \cdot \prod_{i=1}^{k} \int \frac{f}{\mathbb{E}\left(f \mid \mathcal{X}_{i}\right)} d \mu \tag{4}
\end{equation*}
$$

On the other hand, for $1 \leq i \leq k$,

$$
\int \frac{f}{\mathbb{E}\left(f \mid \mathcal{X}_{i}\right)} d \mu=1
$$

and this proves the inequality.

Acknowledgement. The author would like to thank her advisor, Bernard Host, for many helpful discussions and suggestions.

## 2. Proof of Theorem 1.3

In this section, we assume that $\left(Y_{1}, \nu_{1}, S_{1}\right)$ and $\left(Y_{2}, \nu_{2}, S_{2}\right)$ are two ergodic systems, and that $\left(X, \mu, T_{1}, T_{2}\right)=\left(Y_{1} \times Y_{2}, \nu_{1} \times \nu_{2}, S_{1} \times \mathrm{Id}, \mathrm{Id} \times S_{2}\right)$. Then $\left(X, \mu, T_{1}, T_{2}\right)$ is ergodic.

For every $n$, and all bounded functions $\eta, \varphi, \psi$ on $X$, we define

$$
\begin{aligned}
I_{n}(\eta, \varphi, \psi): & =\int \eta \cdot T_{1}^{n} \varphi \cdot T_{2}^{n} \psi d \mu \\
& =\int \eta\left(y_{1}, y_{2}\right) \varphi\left(S_{1}^{n} y_{1}, y_{2}\right) \psi\left(y_{1}, S_{2}^{n} y_{2}\right) d \nu_{1}\left(y_{1}\right) d \nu_{2}\left(y_{2}\right)
\end{aligned}
$$

We prove:
Theorem 2.1. Let $0 \leq f \leq 1$. Then for every $\epsilon>0$, the set

$$
\left\{n \in \mathbb{Z}: \quad I_{n}(f, f, f)>\left(\int f d \mu\right)^{3}-\epsilon\right\}
$$

is syndetic.
Theorem 1.3 follows from Theorem 2.1 with $f=\mathbf{1}_{A}$.
2.1. An ergodic rotation. Let $(Z, \mathcal{Z}, \theta, R)$ denote the common factor of of systems $Y_{1}$ and $Y_{2}$ spanned by the eigenfunctions corresponding to the common eigenvalues of these two systems.

We recall that $Z$ is a compact abelian group, endowed with a Borel $\sigma$-algebra $\mathcal{Z}$ and Haar measure $\theta . R$ is the translation by some fixed $\alpha \in Z$. We consider $\mathcal{Z}$ as a sub- $\sigma$-algebra of both $\left(Y_{1}, \mathcal{Y}_{1}\right)$ and $\left(Y_{2}, \mathcal{Y}_{2}\right)$.

For a bounded function $\eta$ on $X$, we write

$$
\widehat{\eta}:=\mathbb{E}\left(\eta \mid \mathcal{Z} \times \mathcal{Y}_{2}\right) \text { and } \widetilde{\eta}:=\mathbb{E}\left(\eta \mid \mathcal{Y}_{1} \times \mathcal{Z}\right)
$$

We begin with a classical lemma (see for example [10], Lemma 4.18):

## Lemma 2.2.

$$
\mathcal{I}\left(S_{1} \times S_{2}\right) \subset \mathcal{Z} \times \mathcal{Z}
$$

Proof. For $i=1,2$, let $\mathcal{K}_{i}$ be the Kronecker factor of $\left(Y_{i}, S_{i}\right)$. It is known that any $S_{1} \times S_{2}$-invariant function on $Y_{1} \times Y_{2}$ is measurable with respect to $\mathcal{K}_{1} \times \mathcal{K}_{2}$.

Let $\left\{f_{i}\right\}$ (resp. $\left.\left\{g_{j}\right\}\right)$ be an orthonormal basis of $L^{2}\left(\mathcal{K}_{1}, \nu_{1}\right)$ (resp. $\left.L^{2}\left(\mathcal{K}_{2}, \nu_{2}\right)\right)$ consisting of eigenfunctions of $S_{1}$ (resp. $S_{2}$ ), then the set $\left\{f_{i} \otimes \bar{g}_{j}\right\}$ is an orthonormal basis of $L^{2}\left(\mathcal{K}_{1} \times \mathcal{K}_{2}, \nu_{1} \times \nu_{2}\right)$.

Any $F \in L^{\infty}\left(Y_{1} \times Y_{2}\right)$ such that $F$ is invariant under $S_{1} \times S_{2}$ can be written as

$$
F=\sum_{i, j} c_{i, j} f_{i} \otimes \bar{g}_{j} .
$$

for some constants $c_{i, j}$ with $\sum_{i, j}\left|c_{i, j}\right|^{2} \leq \infty$. Write $S_{1} f_{i}=e\left(t_{i}\right) f_{i}$, $S_{2} g_{j}=e\left(s_{j}\right) g_{j}$. By invariance of $F$ under $S_{1} \times S_{2}$, we have

$$
\sum_{i, j} c_{i, j} f_{i} \otimes \bar{g}_{j}=\sum_{i, j} c_{i, j} e\left(t_{i}-s_{j}\right) f_{i} \otimes \bar{g}_{j} .
$$

Hence $c_{i, j} \neq 0$ only when $t_{i}=s_{j}$. This completes the proof.
Lemma 2.3. Let $\eta, \varphi, \psi$ be bounded functions on $X$. Then the averages of the difference $I_{n}(\eta, \varphi, \psi)-I_{n}(\eta, \widehat{\varphi}, \widetilde{\psi})$ converge to 0 .

Proof. Without loss of generality, we can assume that max $\{|\eta|,|\varphi|,|\phi|\} \leq$ 1. By a density argument, we can suppose that $\eta\left(y_{1}, y_{2}\right)=\alpha_{1}\left(y_{1}\right) \beta_{1}\left(y_{2}\right)$, $\varphi\left(y_{1}, y_{2}\right)=\alpha_{2}\left(y_{1}\right) \beta_{2}\left(y_{2}\right)$, and that $\psi\left(y_{1}, y_{2}\right)=\alpha_{3}\left(y_{1}\right) \beta_{3}\left(y_{2}\right)$ for some bounded functions $\alpha_{i}$ on $Y_{1}$ and $\beta_{i}$ on $Y_{2}, i=1,2,3$, all of them being bounded by 1 in absolute value. Using the notation introduced above, we have

$$
\widehat{\varphi}\left(y_{1}, y_{2}\right)=\mathbb{E}\left(\alpha_{2} \mid \mathcal{Z}\right)\left(y_{1}\right) \cdot \beta_{2}\left(y_{2}\right) \text { and } \widetilde{\psi}\left(y_{1}, y_{2}\right)=\alpha_{3}\left(y_{1}\right) \cdot \mathbb{E}\left(\beta_{3} \mid \mathcal{Z}\right)\left(y_{2}\right) .
$$

For every $n$,

$$
\begin{aligned}
& I_{n}(\eta, \varphi, \psi)-I_{n}(\eta, \widehat{\varphi}, \widetilde{\psi})=\int\left(\alpha_{1} \alpha_{3}\right)\left(y_{1}\right) \cdot\left(\beta_{1} \beta_{3}\right)\left(y_{2}\right) \cdot\left(\alpha_{2}\left(S_{1}^{n} y_{1}\right) \cdot \beta_{3}\left(S_{2}^{n} y_{2}\right)\right. \\
&\left.-\mathbb{E}\left(\alpha_{2} \mid \mathcal{Z}\right)\left(S_{1}^{n} y_{1}\right) \cdot \mathbb{E}\left(\beta_{3} \mid \mathcal{Z}\right)\left(S_{2}^{n} y_{2}\right)\right) d \nu\left(y_{1}\right) d \nu\left(y_{2}\right)
\end{aligned}
$$

By the Ergodic Theorem, the averages of this expression converge to

$$
\begin{array}{r}
\mathbb{E}\left(\left(\alpha_{2}-\mathbb{E}\left(\alpha_{2} \mid \mathcal{Z}\right)\right) \otimes \beta_{3} \mid \mathcal{I}\left(S_{1} \times S_{2}\right)\right)+ \\
\mathbb{E}\left(\mathbb{E}\left(\alpha_{2} \mid \mathcal{Z}\right) \otimes\left(\beta_{3}-\mathbb{E}\left(\beta_{3} \mid \mathcal{Z}\right)\right) \mid \mathcal{I}\left(S_{1} \times S_{2}\right)\right)
\end{array}
$$

By Lemma 2.2, this is equal to 0 . This completes the proof.
Lemma 2.4. For every common eigenvalue $e(t)$ of $Y_{1}$ and $Y_{2}$, the averages of the difference

$$
e(n t)\left(I_{n}(f, f, f)-I_{n}(f, \widehat{f}, \widetilde{f})\right)
$$

converge to 0 .
Proof. Let $e(t)$ be a common eigenvalue of $Y_{1}$ and $Y_{2}, h$ be a corresponding eigenfunction of $Y_{1}$ with $|h|=1$. Taking $\eta\left(y_{1}, y_{2}\right)=f\left(y_{1}, y_{2}\right) \bar{h}\left(y_{1}\right)$, $\varphi\left(y_{1}, y_{2}\right)=f\left(y_{1}, y_{2}\right) h\left(y_{1}\right)$, and $\psi\left(y_{1}, y_{2}\right)=f\left(y_{1}, y_{2}\right)$, we have

$$
e(n t) I_{n}(f, f, f)=I_{n}(\eta, \varphi, \psi)
$$

Since $\widehat{\varphi}\left(y_{1}, y_{2}\right)=h\left(y_{1}\right) \widehat{f}\left(y_{1}, y_{2}\right)$ and $\widetilde{\psi}\left(y_{1}, y_{2}\right)=\widetilde{f}\left(y_{1}, y_{2}\right)$, we have

$$
e(n t) I_{n}(f, \widehat{f}, \widetilde{f})=I_{n}(\eta, \widehat{\varphi}, \widetilde{\psi})
$$

By Lemma 2.3, the averages of the difference $I_{n}(\eta, \varphi, \psi)-I_{n}(\eta, \widehat{\varphi}, \widetilde{\psi})$ converge to 0 . This completes the proof.
2.2. A well chosen function $\chi$. In the sequel, we fix $f, \epsilon>0$ as in Theorem 2.1. By continuity of the translation in $L^{2}(Z)$, we can choose an open neighborhood $U$ of 0 in $Z$, such that

$$
\text { for every } s \in U, \int\left|\widehat{f}\left(z+s, y_{2}\right)-\widehat{f}\left(z, y_{2}\right)\right|^{2} d \theta(z) d \nu_{2}\left(y_{2}\right)<\epsilon / 4
$$

and

$$
\text { for every } t \in U, \int\left|\widetilde{f}\left(y_{1}, z+t\right)-\widetilde{f}\left(y_{1}, z\right)\right|^{2} d \nu_{1}\left(y_{1}\right) \theta(z)<\epsilon / 4
$$

We have
(5) $\left|I_{0}(f, \widehat{f}, \widetilde{f})-I_{n}(f, \widehat{f}, \widetilde{f})\right|<\epsilon / 2$ for every $n$ such that $n \alpha \in U$.

We choose a continuous function $\chi$ on $Z$, non-negative, satisfying

$$
\int \chi(z) d \theta(z)=1
$$

and

$$
\chi(z)=0 \text { if } z \notin U .
$$

By Lemma 2.4 and a density argument, we immediately get
Corollary 2.5. The averages of the difference

$$
\chi(n \alpha)\left(I_{n}(f, f, f)-I_{n}(f, \widehat{f}, \widetilde{f})\right)
$$

converge to 0 .
We have

## Lemma 2.6.

(6) $\limsup _{N-M \rightarrow \infty}\left|\frac{1}{N-M} \sum_{n \in[M, N)} \chi(n \alpha)\left(I_{0}(f, \widehat{f}, \tilde{f})-I_{n}(f, f, f)\right)\right| \leq \epsilon / 2$.

Proof. By Lemma 2.5, it suffices to prove
(7) $\limsup _{N-M \rightarrow \infty}\left|\frac{1}{N-M} \sum_{n \in[M-N)} \chi(n \alpha)\left(I_{0}(f, \widehat{f}, \widetilde{f})-I_{n}(f, \widehat{f}, \widetilde{f})\right)\right| \leq \epsilon / 2$.

If $n \alpha \notin U$, then $\chi(n \alpha)\left(I_{0}(f, \widehat{f}, \widetilde{f})-I_{n}(f, \widehat{f}, \widetilde{f})\right)=0$. If $n \alpha \in U$, then by (5),

$$
\chi(n \alpha)\left|I_{0}(f, \widehat{f}, \widetilde{f})-I_{n}(f, \widehat{f}, \tilde{f})\right|<\chi(n \alpha) \cdot \epsilon / 2
$$

By ergodicity of $(Z, \theta, R)$, the sequence $\{n \alpha\}$ is uniformly distributed in $Z$. Therefore the average of $\chi(n \alpha)$ converges to 1 , and this gives (7).
2.3. End of the proof. Applying Lemma 1.6 with $k=2, \mathcal{X}_{1}=\mathcal{Z} \times$ $\mathcal{Y}_{2}$ and $\mathcal{X}_{2}=\mathcal{Y}_{1} \times \mathcal{Z}$, we immediately get the following estimate of $I_{0}(f, \widehat{f}, \widetilde{f})$ :

## Corollary 2.7.

$$
I_{0}(f, \widehat{f}, \widetilde{f}) \geq\left(\int f d \mu\right)^{3}
$$

We can now resume the proof of Theorem 2.1:

Proof of Theorem 2.1. By Corollary 2.7, it suffices to prove that the set $E:=\left\{n: I_{n}(f, f, f)>I_{0}(f, \widehat{f}, \widetilde{f})-\epsilon\right\}$ is syndetic.

Suppose by contradiction that $E$ is not syndetic, then there exists a sequence of intervals $\left[M_{i}, N_{i}\right.$ ) whose lengths tend to infinity such that $I_{n}(f) \leq I_{0}(f, \widehat{f}, \widetilde{f})-\epsilon$ for every $i$ and for every $n \in\left[M_{i}, N_{i}\right)$. Taking averages of $\chi(n \alpha)\left(I_{0}(f, \widehat{f}, \widetilde{f})-I_{n}(f)\right)$ over these intervals, we have a contradiction with Lemma 2.6.

## 3. Preliminaries: Magic systems

In this Section, $\left(X, \mathcal{X}, \mu, T_{1}, T_{2}\right)$ is a system. We review some definitions and properties of [12 needed in sequel, and establish more properties that do not appear there.

The results of this section hold for any number of commuting transformations, but we state and prove them only for two transformations.
3.1. The box measure and the box seminorm. Let $X^{*}=X^{4}$. The points of $X^{*}$ are written as $x^{*}=\left(x_{00}, x_{01}, x_{10}, x_{11}\right)$.
3.1.1. Let $\mu_{1}=\mu \times_{\mathcal{I}\left(T_{1}\right)} \mu$ be the relatively independent square of $\mu$ over $\mathcal{I}\left(T_{1}\right)$. This means that $\mu_{1}$ is the measure on $X^{2}$ characterized by:

For $f_{0}, f_{1} \in L^{\infty}(\mu)$, we have

$$
\int f_{0} \otimes f_{1} d \mu_{1}=\int \mathbb{E}\left(f_{0} \mid \mathcal{I}\left(T_{1}\right)\right) \cdot \mathbb{E}\left(f_{1} \mid \mathcal{I}\left(T_{1}\right)\right) d \mu
$$

Then $\mu_{1}$ is invariant under the transformations $T_{1} \times T_{1}, T_{1} \times$ Id and $T_{2} \times T_{2}$. We define the measure $\mu^{*}$ on $X^{*}=X^{4}$ by

$$
\mu^{*}=\mu_{1} \times_{\mathcal{I}\left(T_{2} \times T_{2}\right)} \mu_{1} .
$$

When needed, we write $\mu_{T_{1}, T_{2}}^{*}$ instead of $\mu^{*}$. Then $\mu^{*}$ is invariant under the transformations $T_{1}^{*}$ and $T_{2}^{*}$ given by

$$
\begin{aligned}
& T_{1}^{*}=T_{1} \times \mathrm{Id} \times T_{1} \times \mathrm{Id}, \\
& T_{2}^{*}=T_{2} \times T_{2} \times \mathrm{Id} \times \mathrm{Id} .
\end{aligned}
$$

The projection $\pi_{00}: x \mapsto x_{00}$ is a factor map from $\left(X^{*}, \mu^{*}, T_{1}^{*}, T_{2}^{*}\right)$ to $\left(X, \mu, T_{1}, T_{2}\right)$.

We remark that by construction,

$$
\begin{equation*}
\mu_{T_{1}, T_{2}}^{*}=\mu_{T_{1}^{1}, T_{2}}^{*}=\mu_{T_{1}, T_{2}^{-1}}^{*} . \tag{8}
\end{equation*}
$$

3.1.2. We recall the definition of the seminorm introduced in [12]:

For every $f \in L^{\infty}(\mu)$ on $X$,

$$
\begin{equation*}
\|f\|_{\mu, T_{1}, T_{2}}=\left(\int f\left(x_{00}\right) f\left(x_{01}\right) f\left(x_{10}\right) f\left(x_{11}\right) d \mu^{*}\left(x^{*}\right)\right)^{1 / 4} \tag{9}
\end{equation*}
$$

We generally omit the subscript $\mu$ when there is no ambiguity. By (8), we have

$$
\|f\|_{T_{1}, T_{2}}=\|f\|_{T_{1}^{-1}, T_{2}},
$$

and by Proposition 2 of [12], we have

$$
\|f\|_{T_{1}, T_{2}}=\|f\|_{T_{2}, T_{1}} .
$$

Here are some properties that we need:
Lemma 3.1. Let $f_{0}, f_{1}, f_{2} \in L^{\infty}(\mu)$ with $\left|f_{i}\right| \leq 1 i=0,1,2$, and for every $n$, we write

$$
I_{n}\left(f_{0}, f_{1}, f_{2}\right):=\int f_{0} \cdot T_{1}^{n} f_{1} \cdot T_{2}^{n} f_{2} d \mu
$$

Then for every $t \in \mathbb{T}$,

$$
\begin{equation*}
\limsup _{N-M \rightarrow \infty}\left|\frac{1}{N-M} \sum_{n \in[M, N)} e(n t) I_{n}\left(f_{0}, f_{1}, f_{2}\right)\right| \leq\left\|f_{0}\right\|_{T_{1}, T_{2}} . \tag{10}
\end{equation*}
$$

For $t=0$, this is Proposition 1 of [12], and the general case follows by the same proof.

Let $f_{0}, f_{1}, f_{2}$ and $I_{n}\left(f_{0}, f_{1}, f_{2}\right)$ be as in Lemma 3.1. For every $n$, we have

$$
I_{n}\left(f_{0}, f_{1}, f_{2}\right)=\int T_{2}^{-n} f_{0} \cdot\left(T_{1} T_{2}^{-1}\right)^{n} f_{1} \cdot f_{2} d \mu
$$

Applying Lemma 3.1 to the pair of transformations $\left(T_{2}, T_{1} T_{2}^{-1}\right)$, we have

$$
\begin{equation*}
\limsup _{N-M \rightarrow \infty}\left|\frac{1}{N-M} \sum_{n \in[M, N)} e(n t) I_{n}\left(f_{0}, f_{1}, f_{2}\right)\right| \leq\left\|f_{1}\right\|_{T_{2}, T_{1} T_{2}^{-1}} . \tag{11}
\end{equation*}
$$

and by the same argument,

$$
\begin{equation*}
\limsup _{N-M \rightarrow \infty}\left|\frac{1}{N-M} \sum_{n \in[M, N)} e(n t) I_{n}\left(f_{0}, f_{1}, f_{2}\right)\right| \leq\left\|f_{2}\right\|_{T_{1}, T_{1} T_{2}^{-1}} . \tag{12}
\end{equation*}
$$

Lemma 3.2 (12], Lemma 1). For every $f \in L^{\infty}(\mu)$,

$$
\begin{aligned}
& \|f\|_{T_{1}, T_{2}}= \\
& \left(\lim _{N_{2} \rightarrow+\infty} \frac{1}{N_{2}} \sum_{n_{2}=0}^{N_{2}-1}\left(\lim _{N_{1} \rightarrow+\infty} \frac{1}{N_{1}} \sum_{n_{1}=0}^{N_{1}-1} \int f \cdot T_{1}^{n_{1}} f \cdot T_{2}^{n_{2}} f \cdot T_{1}^{n_{1}} T_{2}^{n_{2}} f d \mu\right)\right)^{1 / 4}
\end{aligned}
$$

In fact, we can take the simultaneous instead of iterated limit by Theorem 1.1 of ([5]).

By the Hölder Inequality, Lemma 3.2 implies that, for every $f \in$ $L^{\infty}(\mu)$,

$$
\begin{equation*}
\|f\|_{T_{1}, T_{2}} \leq\|f\|_{L^{4}(\mu)} \tag{13}
\end{equation*}
$$

Let $\pi:\left(X, \mu, T_{1}, T_{2}\right) \rightarrow\left(Y, \nu, S_{1}, S_{2}\right)$ be a factor map, and $f \in L^{\infty}(\nu)$, by Lemma 3.2, we have

$$
\begin{equation*}
\|f \circ \pi\|_{\mu, T_{1}, T_{2}}=\|f\|_{\nu, S_{1}, S_{2}} . \tag{14}
\end{equation*}
$$

Moreover, by applying Lemma 3.2 to the measure $\mu$ and to each element of its ergodic decomposition, we get:
Lemma 3.3. If $\mu=\int \mu_{x} d \mu(x)$ is the ergodic decomposition of $\mu$ under $T_{1}$ and $T_{2}$ then, for every $f \in L^{\infty}(\mu)$,

$$
\|f\|_{\mu, T_{1}, T_{2}}^{4}=\int\|f\|_{\mu_{x}, T_{1}, T_{2}}^{4} d \mu(x)
$$

3.1.3. We recall some properties about some $\sigma$-algebra on $X$ introduced in [12]:
Lemma 3.4 ([12], Lemma 4). Let $\mathcal{G}$ be the $\sigma$-algebra on $X$ consisting of sets $B$ such that there exists a subset $A$ of $X^{3}$ satisfying the relation

$$
\begin{equation*}
\mathbf{1}_{B}\left(x_{00}\right)=\mathbf{1}_{A}\left(x_{01}, x_{10}, x_{11}\right) \tag{15}
\end{equation*}
$$

for $\mu^{*}$-almost every $x=\left(x_{00}, x_{01}, x_{10}, x_{11}\right) \in X^{4}$. Then for every $f \in$ $L^{\infty}(\mu)$, we have

$$
\begin{equation*}
\|f\|_{T_{1}, T_{2}}=0 \text { if and only if } \mathbb{E}_{\mu}(f \mid \mathcal{G})=0 . \tag{16}
\end{equation*}
$$

## Lemma 3.5.

$$
\mathcal{G} \supseteq \mathcal{I}\left(T_{1}\right) \vee \mathcal{I}\left(T_{2}\right) .
$$

Proof. Let $f \in L^{\infty}(\mu)$. If $f$ is invariant under $T_{1}$ then, by the construction of the measure $\mu^{*}$, we have $f\left(x_{00}\right)=f\left(x_{10}\right) \mu^{*}$-almost everywhere, and $f$ is measurable with respect to $\mathcal{G}$. If $f$ is invariant under $T_{2}$ then, by a similar reason, $f\left(x_{00}\right)=f\left(x_{10}\right) \mu^{*}$-almost everywhere, and $f$ is measurable with respect to $\mathcal{G}$. This proves the lemma.

Combining Lemma 3.4 and Lemma 3.5, we immediately get:

Corollary 3.6. Let $\left(X, \mu, T_{1}, T_{2}\right)$ be a system, then for every $f \in$ $L^{\infty}(\mu)$, we have

$$
\text { if }\|f\|_{T_{1}, T_{2}}=0 \text { then } \mathbb{E}\left(f \mid \mathcal{I}\left(T_{1}\right) \vee \mathcal{I}\left(T_{2}\right)\right)=0
$$

3.2. Magic systems. We recall the definition of the magic system:

Definition $3.7(\|2\|)$. A system $\left(X, \mu, T_{1}, T_{2}\right)$ is called a magic system if for any $f \in L^{\infty}(\mu)$ satisfying $\mathbb{E}\left(f \mid \mathcal{I}\left(T_{1}\right) \vee \mathcal{I}\left(T_{2}\right)\right)=0$, we have $\|f\|_{T_{1}, T_{2}}=0$.

Corollary 3.8 (of Corollary 3.6). If $\left(X, \mu, T_{1}, T_{2}\right)$ is a magic system, then

$$
\|f\|_{T_{1}, T_{2}}=0 \text { if and only if } \mathbb{E}\left(f \mid \mathcal{I}\left(T_{1}\right) \vee \mathcal{I}\left(T_{2}\right)\right)=0 .
$$

The next Lemma plays an important role in Section 4.1.
Lemma 3.9. Let $p:\left(Y, \nu, S_{1}, S_{2}\right) \rightarrow\left(X, \mu, T_{1}, T_{2}\right)$ be a factor map and assume that $\left(X, \mu, T_{1}, T_{2}\right)$ is magic. Then for every $f \in L^{\infty}(\mu)$,

$$
\left\|f \circ p-\mathbb{E}\left(f \circ p \mid \mathcal{I}\left(S_{1}\right)\right) \vee \mathcal{I}\left(S_{2}\right)\right\|_{S_{1}, S_{2}}=0
$$

Proof. We write $f=f_{1}+f_{2}$, where $f_{1}$ is $\left.\mathcal{I}\left(T_{1}\right)\right) \vee \mathcal{I}\left(T_{2}\right)$-measurable, and $\left.\mathbb{E}\left(f_{2} \mid \mathcal{I}\left(T_{1}\right)\right) \vee \mathcal{I}\left(T_{2}\right)\right)=0$. Since $\left(X, \mu, T_{1}, T_{2}\right)$ is magic, we have $\left\|f_{2}\right\|_{T_{1}, T_{2}}=0$. By (14), $\left\|f_{2} \circ p\right\|_{S_{1}, S_{2}}=0$. By Lemma 3.6, $\mathbb{E}\left(f_{2} \circ p \mid\right.$ $\left.\mathcal{I}\left(S_{1}\right) \vee \mathcal{I}\left(S_{2}\right)\right)=0$. On the other hand, $f_{1} \circ p$ is measurable with respect to $\mathcal{I}\left(S_{1}\right) \vee \mathcal{I}\left(S_{2}\right)$. Therefore $\mathbb{E}\left(f \circ p \mid \mathcal{I}\left(S_{1}\right) \vee \mathcal{I}\left(S_{2}\right)\right)=f_{1} \circ p$ and $\left\|f \circ p-\mathbb{E}\left(f \circ p \mid \mathcal{I}\left(S_{1}\right) \vee \mathcal{I}\left(S_{2}\right)\right)\right\|_{S_{1}, S_{2}}=\left\|f_{2} \circ p\right\|_{S_{1}, S_{2}}=0$.

One of the main results of [12] is the following theorem:
Theorem 3.10 (Theorem 1, [12]). The system $\left(X^{*}, \mu^{*}, T_{1}^{*}, T_{2}^{*}\right)$ built in Section 3.1 is magic.

Since $\pi_{00}:\left(X^{*}, \mu^{*}, T_{1}^{*}, T_{2}^{*}\right) \rightarrow\left(X, \mu, T_{1}, T_{2}\right)$ is a factor map, we have:
Corollary 3.11. Every system ( $X, \mu, T_{1}, T_{2}$ ) admits a magic extension.
We can not use this result directly because the system $X^{*}$ is not ergodic even if $X$ is ergodic.
3.3. The existence of ergodic magic extension. In this Section we show:

Theorem 3.12. Every ergodic system $\left(X, \mathcal{X}, \mu, T_{1}, T_{2}\right)$ admits an ergodic magic extension.

Recall that we assume that $(X, \mathcal{X}, \mu)$ is a standard probability space. Then $X$ admits a dense countably generated $\sigma$-algebra $\mathcal{X}^{\prime}$. Substituting this algebra for $\mathcal{X}$, we reduce to the case that the $\sigma$-algebra $\mathcal{X}$ is countably generated. In the sequel,

$$
\mu=\int \mu_{x} d \mu(x)
$$

denotes the ergodic decomposition of $\mu$ under $T_{1}$ and $T_{2}$.
Theorem 3.12 follows from the following proposition:
Proposition 3.13. Let $\left(X, \mathcal{X}, \mu, T_{1}, T_{2}\right)$ be a magic system. Then for $\mu$-almost every $x$, the system $\left(X, \mathcal{X}, \mu_{x}, T_{1}, T_{2}\right)$ is magic.
Proof of Theorem 3.12 assuming Proposition 3.13. By Corollary 3.11, there exist a magic system $\left(X^{*}, \mu^{*}, T_{1}^{*}, T_{2}^{*}\right)$ and a factor map $\pi: X^{*} \rightarrow$ $X$. Let $\mu^{*}=\int \mu_{x}^{*} d \mu^{*}(x)$ be the ergodic decomposition of $\mu^{*}$ under $T_{1}^{*}$ and $T_{2}^{*}$. Then $\mu=\int \pi\left(\mu_{x}^{*}\right) d \mu^{*}(x)$. Since the measures $\pi\left(\mu_{x}^{*}\right)$ are invariant under $T_{1}$ and $T_{2}$ and since $\mu$ is ergodic, then for $\mu^{*}$-almost every $x$ we have $\pi\left(\mu_{x}^{*}\right)=\mu$, and thus $\pi:\left(X^{*}, \mu_{x}^{*}, T_{1}^{*}, T_{2}^{*}\right) \rightarrow\left(X, \mu, T_{1}, T_{2}\right)$ is a factor map.

Before giving the proof of Proposition 3.13, we need the following two lemmas. We begin with the notation. We denote

$$
\mathcal{A}=\mathcal{I}\left(T_{1}\right) \wedge \mathcal{I}\left(T_{2}\right) \text { and } \mathcal{W}=\mathcal{I}\left(T_{1}\right) \vee \mathcal{I}\left(T_{2}\right) .
$$

We recall that for $\mu$-almost every $x$, the measure $\mu_{x}$ is invariant and ergodic under $T_{1}$ and $T_{2}$. The map $x \mapsto \mu_{x}$ is $\mathcal{A}$-measurable and for every bounded function $f$ on $X$,

$$
\mathbb{E}_{\mu}(f \mid \mathcal{A})(x)=\int f d \mu_{x} \text { for } \mu \text {-almost every } x
$$

Lemma 3.14. There exists a countable family $\Phi$ of bounded $\mathcal{W}$-measurable functions on $X$, which is dense in $L^{p}(\mathcal{W}, \nu)$ for every $p \in[1,+\infty)$ and every probability measure $\nu$ on $(X, \mathcal{X})$ which is invariant under $T_{1}$ and $T_{2}$.
Proof. Let $\mathcal{F}=\left\{f_{n}: n \in \mathbb{N}\right\}$ be a countable family of bounded $\mathcal{X}$ measurable functions on $X$ which is dense in $L^{p}(\nu)$ for every $p \in$ $[1,+\infty)$ and every probability measure $\nu$ on $(X, \mathcal{X})$. For each $n$, define

$$
g_{n}(x)= \begin{cases}\lim _{K \rightarrow \infty} \frac{1}{K} \sum_{k=0}^{K-1} f_{n}\left(T_{1}^{k} x\right) & \text { if this limit exists; } \\ 0 & \text { otherwise }\end{cases}
$$

Then each function $g_{n}$ is bounded and invariant under $T_{1}$. For every $n$, let $h_{n}$ be the function associated to $T_{2}$ by the same construction.

Let $\mathcal{F}_{1}=\left\{g_{n}: n \in \mathbb{N}\right\}$ and $\mathcal{F}_{2}=\left\{h_{n}: n \in \mathbb{N}\right\}$. Then, for $i=$ $1,2, \mathcal{F}_{i}$ is dense in $L^{p}\left(\mathcal{I}\left(T_{i}\right), \nu\right)$ for every $p \in[1,+\infty)$ and every $T_{i^{-}}$ invariant probability measure $\nu$ on $(X, \mathcal{X})$. Let $\Phi$ be family of functions consisting in finite sums of functions of the form $g h$ with $g \in \mathcal{F}_{1}$ and $h \in \mathcal{F}_{2}$, then $\Phi$ satisfies the lemma.
Lemma 3.15. Let $f$ be a bounded function on $X$ and let $\tilde{f}$ be a $\mathcal{W}$ measurable representative of $\mathbb{E}_{\mu}(f \mid \mathcal{W})$. Then, for $\mu$-almost every $x$, we have

$$
\mathbb{E}_{\mu_{x}}(f \mid \mathcal{W})=\widetilde{f} \quad \mu_{x} \text {-almost everywhere. }
$$

Proof. Let $\Phi=\left\{\phi_{n}: n \in \mathbb{N}\right\}$ be the family of functions given by Lemma 3.14. For every $n, \mathbb{E}_{\mu}\left((f-\widetilde{f}) \phi_{n} \mid \mathcal{W}\right)=0$ and, since $\mathcal{A} \subset \mathcal{W}$, $\mathbb{E}_{\mu}\left((f-\widetilde{f}) \phi_{n} \mid \mathcal{A}\right)=0 \mu$-almost everywhere. That is, $\int(f-\widetilde{f}) \phi_{n} d \mu_{x}=$ 0 for $\mu$-almost every $x$. Thus there exists a set $A_{n}$ with $\mu\left(A_{n}\right)=1$ such that $\int(f-\widetilde{f}) \phi_{n} d \mu_{x}=0$ for every $x \in A_{n}$.

Let $A$ be the intersection of the sets $A_{n}$ for $n \in \mathbb{N}$. Then $\mu(A)=1$. Let $x \in A$. We have $\int(f-\widetilde{f}) \phi_{n} d \mu_{x}=0$ for every $n$ and, by density of the family $\left\{\phi_{n}: n \in \mathbb{N}\right\}$ in $L^{1}\left(\mathcal{W}, \mu_{x}\right)$, we have $\mathbb{E}_{\mu_{x}}(f-\widetilde{f} \mid \mathcal{W})=0$. Since $\widetilde{f}$ is measurable with respect to $\mathcal{W}$, we have $\mathbb{E}_{\mu_{x}}(f \mid \mathcal{W})=\widetilde{f}$.
Proof of Proposition 3.13. Let $\mathcal{F}=\left\{f_{n}: n \in \mathbb{N}\right\}$ be a countable family of bounded $\mathcal{X}$-measurable functions on $X$ which is dense in $L^{p}(\nu)$ for every $p \in[1,+\infty)$ and every probability measure $\nu$ on $(X, \mathcal{X})$.

For each $n$, let $\widetilde{f}_{n}$ be a $\mathcal{W}$-measurable representative of $\mathbb{E}_{\mu}\left(f_{n} \mid \mathcal{W}\right)$. By Lemma 3.15, for every $n$ we have that $\widetilde{f}_{n}=\mathbb{E}_{\mu_{x}}\left(f_{n} \mid \mathcal{W}\right) \mu_{x}$-almost everywhere for $\mu$-almost every $x$, and there exists a set $A_{n}$ such that $\mu\left(A_{n}\right)=1$ and,

$$
\text { for every } x \in A_{n}, \quad \mathbb{E}_{\mu_{x}}\left(f_{n} \mid \mathcal{W}\right)=\widetilde{f}_{n} \quad \mu_{x} \text {-a.e. }
$$

Fix an integer $n \in \mathbb{N}$. Since $\mathbb{E}_{\mu}\left(f_{n}-\widetilde{f}_{n} \mid \mathcal{W}\right)=0$ and $\left(X, \mu, T_{1}, T_{2}\right)$ is magic, we have $\left\|f_{n}-\widetilde{f}_{n}\right\|_{\mu, T_{1}, T_{2}}=0$.

By Lemma 3.3, we have

$$
0=\left\|f_{n}-\widetilde{f}_{n}\right\|_{\mu, T_{1}, T_{2}}^{4}=\int\left\|f_{n}-\widetilde{f}_{n}\right\|_{\mu_{x}, T_{1}, T_{2}}^{4} d \mu(x) .
$$

So $\left\|\left\|f_{n}-\widetilde{f}_{n}\right\|_{\mu_{x}, T_{1}, T_{2}}=0\right.$ for $\mu$-almost every $x$. There exists a set $B_{n} \subset X$ such that $\mu\left(B_{n}\right)=1$ and

$$
\text { for every } x \in B_{n}, \quad\left\|f_{n}-\widetilde{f}_{n}\right\|_{\mu_{x}, T_{1}, T_{2}}=0 .
$$

Define

$$
C=\bigcap_{n \in N}\left(A_{n} \cap B_{n}\right) .
$$

We have $\mu(C)=1$. We check that for every $x \in C$ the system ( $X, \mu_{x}, T_{1}, T_{2}$ ) is magic.

Let $x \in C$ and $f$ be a bounded measurable function on $X$, we have to show that $\left\|f-\mathbb{E}_{\mu_{x}}(f \mid \mathcal{W})\right\|_{\mu_{x}, T_{1}, T_{2}}=0$.

Since the family $\left\{f_{n}\right\}$ is dense in $L^{4}\left(\mu_{x}\right)$, there exists a sequence $\left(n_{i}\right)$ of integers such that $f_{n_{i}}$ converges to $f$ in $L^{4}\left(\mu_{x}\right)$. By (13) applied to the measure $\mu_{x}$, we have

$$
\begin{equation*}
\left\|f_{n_{i}}-f\right\|_{\mu_{x}, T_{1}, T_{2}} \rightarrow 0 . \tag{17}
\end{equation*}
$$

On the other hand, since $f_{n_{i}}$ converges to $f$ in $L^{4}\left(\mu_{x}\right)$, we have that $\mathbb{E}_{\mu_{x}}\left(f_{n_{i}} \mid \mathcal{W}\right)$ converges to $\mathbb{E}_{\mu_{x}}(f \mid \mathcal{W})$ in $L^{4}\left(\mu_{x}\right)$ and, as above,

$$
\begin{equation*}
\left\|\mathbb{E}_{\mu_{x}}\left(f_{n_{i}} \mid \mathcal{W}\right)-\mathbb{E}_{\mu_{x}}(f \mid \mathcal{W})\right\|_{\mu_{x}, T_{1}, T_{2}} \rightarrow 0 \tag{18}
\end{equation*}
$$

For every $i$, since $x \in A_{n_{i}}$ we have that $\mathbb{E}_{\mu_{x}}\left(f_{n_{i}} \mid \mathcal{W}\right)=\widetilde{f}_{n_{i}}$ ( $\mu_{x^{-}}$-a.e.). Since $x \in B_{n_{i}}$, we have $\left\|f_{n_{i}}-\widetilde{f}_{n_{i}}\right\| \|_{\mu_{x}, T_{1}, T_{2}}=0$ and thus

$$
\begin{equation*}
\left\|f_{n_{i}}-\mathbb{E}_{\mu_{x}}\left(f_{n_{i}} \mid \mathcal{W}\right)\right\|_{\mu_{x}, T_{1}, T_{2}}=0 \tag{19}
\end{equation*}
$$

For every $i$,

$$
\begin{aligned}
& \left\|f-\mathbb{E}_{\mu_{x}}(f \mid \mathcal{W})\right\|_{\mu_{x}, T_{1}, T_{2}} \\
& \leq\left\|f_{n_{i}}-f\right\|_{\mu_{x}, T_{1}, T_{2}}+\left\|\mathbb{E}_{\mu_{x}}\left(f_{n_{i}} \mid \mathcal{W}\right)-\mathbb{E}_{\mu_{x}}(f \mid \mathcal{W})\right\|_{\mu_{x}, T_{1}, T_{2}} \\
& \\
& \quad+\left\|f_{n_{i}}-\mathbb{E}_{\mu_{x}}\left(f_{n_{i}} \mid \mathcal{W}\right)\right\|_{\mu_{x}, T_{1}, T_{2}} .
\end{aligned}
$$

Combining (17), (18), (19), we get $\left\|f-\mathbb{E}_{\mu_{x}}(f \mid \mathcal{W})\right\|_{\mu_{x}, T_{1}, T_{2}}=0$ for every $x \in C$. This finishes the proof.

## 4. Proof of Theorem 1.1

In this section, we prove:
Theorem 4.1. Let $\left(X, \mu, T_{1}, T_{2}\right)$ be an ergodic system and $0 \leq f \leq 1$. Then for every $\epsilon>0$, the set

$$
\left\{n \in \mathbb{Z}: \quad \int f \cdot T_{1}^{n} f \cdot T_{2}^{n} f d \mu>\left(\int f d \mu\right)^{4}-\epsilon\right\}
$$

is syndetic.
Theorem 1.1 follows from Theorem 4.1 with $f=\mathbf{1}_{A}$.
In this section, we fix the ergodic system $\left(X, \mu, T_{1}, T_{2}\right)$, and the function $f$ as in Theorem 4.1, and for every $n$, we write

$$
I_{n}:=\int f \cdot T_{1}^{n} f \cdot T_{2}^{n} f d \mu
$$

4.1. Construction of ergodic magic extensions. We write $T_{3}=$ $T_{1} T_{2}^{-1}$. We use Theorem 3.12 three times:

Let $\left(X^{\prime}, \mu^{\prime}, T_{1}^{\prime}, T_{2}^{\prime}\right)$ be an ergodic magic extension of ( $X, \mu, T_{1}, T_{2}$ ) and let $\pi^{\prime}: X^{\prime} \rightarrow X$ be the factor map. Write $T_{3}^{\prime}=T_{1}^{\prime} T_{2}^{\prime-1}$. Then the system ( $X^{\prime}, \mu^{\prime}, T_{1}^{\prime}, T_{3}^{\prime}$ ) is ergodic.

Let ( $X^{\prime \prime}, \mu^{\prime \prime}, T_{1}^{\prime \prime}, T_{3}^{\prime \prime}$ ) be an ergodic magic extension of ( $X^{\prime}, \mu^{\prime}, T_{1}^{\prime}, T_{3}^{\prime}$ ) and let $\pi^{\prime \prime}: X^{\prime \prime} \rightarrow X^{\prime}$ be the factor map. Write $T_{2}^{\prime \prime}=T_{1}^{\prime \prime} T_{3}^{\prime \prime-1}$. Then the system $\left(X^{\prime \prime}, \mu^{\prime \prime}, T_{2}^{\prime \prime}, T_{3}^{\prime \prime}\right)$ is ergodic.

Let $\left(\bar{X}, \bar{\mu}, \bar{T}_{2}, \bar{T}_{3}\right)$ be an ergodic magic extension of ( $X^{\prime \prime}, \mu^{\prime \prime}, T_{2}^{\prime \prime}, T_{3}^{\prime \prime}$ ) and let $\bar{\pi}: \bar{X} \rightarrow X^{\prime \prime}$ be the factor map. Write $\bar{T}_{1}=\bar{T}_{2} \bar{T}_{3}$. Then the three systems $\left(\bar{X}, \bar{\mu}, \bar{T}_{2}, \bar{T}_{3}\right),\left(\bar{X}, \bar{\mu}, \bar{T}_{1}, \bar{T}_{3}\right)$ and $\left(\bar{X}, \bar{\mu}, \bar{T}_{1}, \bar{T}_{2}\right)$ are all ergodic, and $\pi^{\prime} \circ \pi^{\prime \prime} \circ \bar{\pi}$ is a factor map from $\left.\left(\bar{X}, \bar{\mu}, \bar{T}_{1}, \bar{T}_{2}, \bar{T}_{3}\right)\right)$ to $\left(X, \mu, T_{1}, T_{2}, T_{3}\right)$.

In the sequel, we write

$$
\begin{align*}
h & =: f \circ \pi^{\prime} \circ \pi^{\prime \prime} \circ \bar{\pi} \\
g_{0} & =: \mathbb{E}\left(h \mid \mathcal{I}\left(\bar{T}_{1}\right) \vee \mathcal{I}\left(\bar{T}_{2}\right)\right), \\
g_{1} & =: \mathbb{E}\left(h \mid \mathcal{I}\left(\bar{T}_{1}\right) \vee \mathcal{I}\left(\bar{T}_{3}\right)\right),  \tag{20}\\
g_{2} & =: \mathbb{E}\left(h \mid \mathcal{I}\left(\bar{T}_{2}\right) \vee \mathcal{I}\left(\bar{T}_{3}\right)\right) .
\end{align*}
$$

For every $n$, and all bounded functions $\eta, \varphi, \psi$ on $\bar{X}$, we define

$$
\bar{I}_{n}(\eta, \varphi, \psi):=\int \eta \cdot \bar{T}_{1}^{n} \varphi \cdot \bar{T}_{2}^{n} \psi d \bar{\mu} .
$$

In particular,

$$
\begin{equation*}
\bar{I}_{n}(h, h, h)=I_{n} . \tag{21}
\end{equation*}
$$

We write also

$$
J_{n}:=\bar{I}_{n}\left(g_{0}, g_{1}, g_{2}\right) .
$$

We have
Proposition 4.2. Let $t \in \mathbb{T}$. Then the averages of the difference $e(n t)\left(I_{n}-J_{n}\right)$ converge to 0 .

Proof. Since $\left(\bar{X}, \bar{\mu}, \bar{T}_{2}, \bar{T}_{3}\right)$ is magic, by Definition 3.7 of magic systems and the definition of $g_{2}$,

$$
\left\|h-g_{2}\right\|_{\bar{T}_{2}, \bar{T}_{3}}=0 .
$$

Then by Lemma 3.1, the averages of the difference of

$$
e(n t)\left(\bar{I}_{n}(h, h, h)-\bar{I}_{n}\left(h, h, g_{2}\right)\right)
$$

converge to 0 . Since $\left(X^{\prime \prime}, \mu^{\prime \prime}, T_{1}^{\prime \prime}, T_{3}^{\prime \prime}\right)$ is magic, by Lemma 3.9,

$$
\left\|h-g_{1}\right\|_{\bar{T}_{1}, \bar{T}_{3}}=0 .
$$

Then by Lemma 3.1, the averages of the difference of

$$
e(n t)\left(\bar{I}_{n}\left(h, h, g_{2}\right)-\bar{I}_{n}\left(h, g_{1}, g_{2}\right)\right)
$$

converge to 0 . Since $\left(X^{\prime}, \mu^{\prime}, T_{1}^{\prime}, T_{2}^{\prime}\right)$ is magic, by the same argument as above, the averages of the difference of

$$
e(n t)\left(\bar{I}_{n}\left(h, g_{1}, g_{2}\right)-J_{n}\right)
$$

converge to 0 . Summing up all the results above, we get the announced result.

In the sequel, we write

$$
\overline{\mathcal{M}}:=\mathcal{I}\left(\bar{T}_{1}\right) \vee \mathcal{I}\left(\bar{T}_{2}\right) \vee \mathcal{I}\left(\bar{T}_{3}\right),
$$

and

$$
g:=\mathbb{E}(h \mid \overline{\mathcal{M}})
$$

We remark that

$$
\begin{aligned}
& \mathbb{E}\left(g \mid \mathcal{I}\left(\bar{T}_{1}\right) \vee \mathcal{I}\left(\bar{T}_{2}\right)\right)=g_{0}, \\
& \mathbb{E}\left(g \mid \mathcal{I}\left(\bar{T}_{1}\right) \vee \mathcal{I}\left(\bar{T}_{3}\right)\right)=g_{1}, \\
& \mathbb{E}\left(g \mid \mathcal{I}\left(\bar{T}_{2}\right) \vee \mathcal{I}\left(\bar{T}_{3}\right)\right)=g_{2} .
\end{aligned}
$$

Furthermore

$$
\begin{equation*}
\int g d \bar{\mu}=\int h d \bar{\mu}=\int f d \mu \tag{22}
\end{equation*}
$$

4.2. Reduction to a special system. For $i=1,2,3$, let $\left(Y_{i}, \mathcal{Y}_{i}, \nu_{i}\right)$ be a factor of $\bar{X}$ which is isomorphic to $\left(\bar{X}, \mathcal{I}\left(\bar{T}_{i}\right), \bar{\mu}\right)$. Denote by $\pi_{i}$ : $\bar{X} \rightarrow Y_{i}$ the factor map. Then $\nu_{i}:=\pi_{i}(\bar{\mu})$.

Since for $j=1,2,3$, every $\sigma$-algebra $\mathcal{I}\left(\bar{T}_{i}\right)$ is invariant under the transformation $\bar{T}_{j}$, then each of these transformations induces a measure preserving transformation on each factor $Y_{i}$.

The transformation $\bar{T}_{1}$ induces the identity on $Y_{1}$ and the transformations $\bar{T}_{2}$ and $\bar{T}_{3}$ induce the same transformation on $Y_{1}$, which we write $R_{1}$. Therefore we have

$$
R_{1} \circ \pi_{1}=\pi_{1} \circ \bar{T}_{2}=\pi_{1} \circ \bar{T}_{3}
$$

By ergodicity of ( $\bar{X}, \overline{\mathcal{X}}, \bar{\mu}, \bar{T}_{1}, \bar{T}_{2}$ ), the system $\left(\bar{X}, \mathcal{I}\left(\bar{T}_{1}\right), \bar{\mu}, \bar{T}_{2}\right)$ is ergodic and thus ( $Y_{1}, \nu_{1}, R_{1}$ ) is ergodic.

In the same way, we get a measure preserving transformation $R_{2}$ on $Y_{2}$ and a measure preserving transformation $R_{3}$ on $Y_{3}$ with

$$
\begin{gathered}
R_{2} \circ \pi_{2}=\pi_{2} \circ \bar{T}_{1}=\pi_{2} \circ \bar{T}_{3}^{-1} \\
R_{3} \circ \pi_{3}=\pi_{3} \circ \bar{T}_{1}=\pi_{3} \circ \bar{T}_{2}
\end{gathered}
$$

As above, we have $\left(Y_{2}, \nu_{2}, R_{2}\right)$ and $\left(Y_{3}, \nu_{3}, R_{3}\right)$ are ergodic.

Let $Y=Y_{1} \times Y_{2} \times Y_{3}$ and $\mathcal{Y}$ be the product $\sigma$-algebra on $Y$. Define $\pi:(\bar{X}, \overline{\mathcal{M}}) \rightarrow(Y, \mathcal{Y})$ by $\pi=\pi_{1} \times \pi_{2} \times \pi_{3}$. Let $\nu$ be the image of $\mu$ under $\pi$. Define

$$
\begin{align*}
& S_{1}=\mathrm{Id} \times R_{2} \times R_{3}, \\
& S_{2}=R_{1} \times \mathrm{Id} \times R_{3},  \tag{23}\\
& S_{3}=R_{1} \times R_{2}^{-1} \times \mathrm{Id} .
\end{align*}
$$

We have
Lemma 4.3. $\pi$ is an isomorphism from $\left(\bar{X}, \overline{\mathcal{M}}, \bar{\mu}, \bar{T}_{1}, \bar{T}_{2}, \bar{T}_{3}\right)$ to $\left(Y, \mathcal{Y}, \nu, S_{1}, S_{2}, S_{3}\right)$.

We can thus identify these two systems and consider $g$ as a function on $Y$. Then we have

$$
\begin{aligned}
g_{0} & =\mathbb{E}\left(g \mid Y_{1} \times Y_{2}\right), \\
g_{1} & =\mathbb{E}\left(g \mid Y_{1} \times Y_{3}\right), \\
g_{2} & =\mathbb{E}\left(g \mid Y_{2} \times Y_{3}\right) .
\end{aligned}
$$

By (22), we have

$$
\begin{equation*}
\int g d \nu=\int f d \mu \tag{24}
\end{equation*}
$$

4.3. Description of the special system. We recall the properties that will be used in the sequel:
(a) For $i=1,2,3,\left(Y_{i}, \mathcal{Y}_{i}, \nu_{i}, R_{i}\right)$ is an ergodic system.
(b) $Y=Y_{1} \times Y_{2} \times Y_{3}$ is endowed with the product $\sigma$-algebra $\mathcal{Y}$. The projection $Y \rightarrow Y_{i}$ is written $\pi_{i}$.
(c) $\nu$ is a joining of $\nu_{1}, \nu_{2}$ and $\nu_{3}$, meaning that, for $i=1,2,3$, the projection of $\nu$ on $Y_{i}$ is equal to $\nu_{i}$. Moreover $\nu$ is invariant under the transformations $S_{1}, S_{2}, S_{3}$ defined in (23).
(d) For $i=1,2,3, \mathcal{Y}_{i}=\mathcal{I}\left(S_{i}\right)$.
(e) $\left(Y, \mathcal{Y}, \nu, S_{1}, S_{2}\right)$ is ergodic.

Remark. For $i \neq j, Y$ is ergodic under $S_{i}$ and $S_{j}$. By (d), the $\sigma$ algebra $\mathcal{Y}_{i}$ and $\mathcal{Y}_{j}$ are independent. In other words, the projection of $\nu$ on $Y_{i} \times Y_{j}$ is equal to $\nu_{i} \times \nu_{j}$.

The following proposition is a more precise description of $\nu$ :
Proposition 4.4. Let $(Z, \theta, R)$ be the common factor of the systems $Y_{1}, Y_{2}, Y_{3}$ spanned by the eigenfunctions corresponding to the common
eigenvalues of these three systems. Then for any $\alpha_{1} \in L^{\infty}\left(Y_{1}\right), \alpha_{2} \in$ $L^{\infty}\left(Y_{2}\right)$ and $\alpha_{3} \in L^{\infty}\left(Y_{3}\right)$, we have
$\int \alpha_{1}\left(y_{1}\right) \alpha_{2}\left(y_{2}\right) \alpha_{3}\left(y_{3}\right) d \nu=\int \mathbb{E}\left(\alpha_{1} \mid Z\right)\left(z_{1}\right) \mathbb{E}\left(\alpha_{2} \mid Z\right)\left(z_{2}\right) \mathbb{E}\left(\alpha_{3} \mid Z\right)\left(z_{3}\right) d m$
where $m$ is the image of $\nu$ on $Z \times Z \times Z$. In other words, $\nu$ is the conditionally independent product measure over $Z \times Z \times Z$.

Proof. For $i=1,2,3$ let $\mathcal{K}_{i}$ be the Kronecker factor of $Y_{i}$. Since $\nu$ is invariant under $S_{1}$, we have

$$
\int \alpha_{1}\left(y_{1}\right) \alpha_{2}\left(y_{2}\right) \alpha_{3}\left(y_{3}\right) d \nu=\frac{1}{N} \sum_{n=1}^{N} \int \alpha_{1}\left(y_{1}\right) \cdot R_{2}^{n} \alpha_{2}\left(y_{2}\right) \cdot R_{3}^{n} \alpha_{3}\left(y_{3}\right) d \nu
$$

for every $N$. Since $Y_{2}, Y_{3}$ are independent, taking the limit as $N \rightarrow \infty$, we have

$$
\int \alpha_{1}\left(y_{1}\right) \alpha_{2}\left(y_{2}\right) \alpha_{3}\left(y_{3}\right) d \nu=\int \alpha_{1}\left(y_{1}\right) \cdot \mathbb{E}_{\nu_{2} \times \nu_{3}}\left(\alpha_{2} \otimes \alpha_{3} \mid \mathcal{I}\left(R_{2} \times R_{3}\right)\right)\left(y_{2}, y_{3}\right) d \nu
$$

Note that the functions $\alpha_{2} \otimes \alpha_{3}$ and $\mathbb{E}\left(\alpha_{2} \mid \mathcal{K}_{2}\right) \otimes \mathbb{E}\left(\alpha_{3} \mid \mathcal{K}_{3}\right)$ have the same conditional expectation on $\mathcal{I}_{\nu_{2} \times \nu_{3}}\left(R_{2} \times R_{3}\right)$. Therefore we have

$$
\int \alpha_{1}\left(y_{1}\right) \alpha_{2}\left(y_{2}\right) \alpha_{3}\left(y_{3}\right) d \nu=\int \alpha_{1}\left(y_{1}\right) \mathbb{E}\left(\alpha_{2} \mid \mathcal{K}_{2}\right)\left(y_{2}\right) \mathbb{E}\left(\alpha_{3} \mid \mathcal{K}_{3}\right)\left(y_{3}\right) d \nu
$$

By the same computation, substituting $S_{2}$ for $S_{1}$, we obtain

$$
\begin{equation*}
\int \alpha_{1}\left(y_{1}\right) \alpha_{2}\left(y_{2}\right) \alpha_{3}\left(y_{3}\right) d \nu=\int \mathbb{E}\left(\alpha_{1} \mid \mathcal{K}_{1}\right)\left(y_{1}\right) \mathbb{E}\left(\alpha_{2} \mid \mathcal{K}_{2}\right)\left(y_{2}\right) \mathbb{E}\left(\alpha_{3} \mid \mathcal{K}_{3}\right)\left(y_{3}\right) d \nu \tag{26}
\end{equation*}
$$

On the other hand, for $i=1,2,3$, the right hand side of (25) remains unchanged if $\mathbb{E}\left(\alpha_{i} \mid \mathcal{K}_{i}\right)$ is substituted for $\alpha_{i}$. Therefore we can reduce to the case that $\alpha_{i}$ is measurable with respect to $\mathcal{K}_{i}$. By density, we can suppose that $\alpha_{i}$ is an eigenfunction of $Y_{i}$ corresponding to the eigenvalue $\lambda_{i}$. By invariance of $\nu$ under $S_{1}$ again, we get

$$
\int \alpha_{1}\left(y_{1}\right) \alpha_{2}\left(y_{2}\right) \alpha_{3}\left(y_{3}\right) d \nu=e\left(\lambda_{2}+\lambda_{3}\right) \int \alpha_{1}\left(y_{1}\right) \alpha_{2}\left(y_{2}\right) \alpha_{3}\left(y_{3}\right) d \nu
$$

thus $\int \alpha_{1}\left(y_{1}\right) \alpha_{2}\left(y_{2}\right) \alpha_{3}\left(y_{3}\right) d \nu=0$ except if $\lambda_{2}=-\lambda_{3}$.
By the same argument, we have that $\int \alpha_{1}\left(y_{1}\right) \alpha_{2}\left(y_{2}\right) \alpha_{3}\left(y_{3}\right) d \nu=0$ except if $\lambda_{1}=\lambda_{2}=-\lambda_{3}$. The announced conclusion follows.

We have the following description of the measure $m$ :

Proposition 4.5. The measure $m$ is the Haar measure on the compact subgroup $H$ of $Z \times Z \times Z$, where

$$
H=\left\{\left(z_{1}, z_{2}, z_{3}\right), z_{1}+z_{2}-z_{3}=0\right\}
$$

Proof. Let $\alpha \in Z$ be the element defining the translation $R$. Since $\nu$ is invariant and ergodic under the transformations $S_{1}$ and $S_{2}$, then the measure $m$ is invariant and ergodic under the transformations Id $\times R \times R$ and $R \times \operatorname{Id} \times R$, that is under the translations by ( $0, \alpha, \alpha$ ) and $(\alpha, 0, \alpha)$. The function $z_{1}+z_{2}-z_{3}$ is invariant under these translations and by ergodicity it is equal to a constant $c$.

Since the map $z \mapsto z-c$ is an isomorphism from $(Z, \theta, R)$ to itself, we can thus reduce to the case that $c=0$ and $m$ is concentrated on $H$. Since the subset $\mathbb{Z} \alpha$ is dense in $Z, m$ is invariant under the translations by $(0, z, z)$ and $\left(z^{\prime}, 0, z^{\prime}\right)$ for any $z, z^{\prime} \in Z$. Since these elements span $H, m$ is invariant under translation by any element of $H$, and thus $m$ is the Haar measure on $H$.
4.4. Study of $J_{n}$. In this section, let $\left(Y, \nu, S_{1}, S_{2}\right)$ be as described in Section 4.3. We keep using the notation $g, g_{0}, g_{1}, g_{2}$ and $J_{n}=$ $\bar{I}_{n}\left(g_{0}, g_{1}, g_{2}\right)$ introduced above. From this point, the proof is parallel to that in Section 2.2.

Proposition 4.6. For every $\epsilon>0$, there exists an open neighborhood $U$ of 0 in $Z$, such that

$$
\left|J_{0}-J_{n}\right|<\epsilon .
$$

for every $n$ such that $n \alpha \in U$.
Proof. More generally, we show for all bounded functions $h_{0}$ on $Y_{1} \times$ $Y_{2} Y, h_{1}$ on $Y_{1} \times Y_{3}$, and $h_{2}$ on $Y_{2} \times Y_{3}$, that there exists an open neighborhood $U$ of 0 in $Z$, such that

$$
\left|\int h_{0} \cdot h_{1} \cdot h_{2} d \nu-\int h_{0} \cdot S_{1}^{n} h_{1} \cdot S_{2}^{n} h_{2} d \nu\right|<\epsilon
$$

for every $n$ such that $n \alpha \in U$.
Without loss of generality, we can suppose that max $\left\{\left|h_{0}\right|,\left|h_{1}\right|,\left|h_{2}\right|\right\} \leq$ 1. We first suppose that $h_{0}, h_{1}, h_{2}$ are product functions: $h_{0}\left(y_{1}, y_{2}\right)=$ $\alpha\left(y_{1}\right) \beta\left(y_{2}\right), h_{1}\left(y_{1}, y_{3}\right)=\alpha^{\prime}\left(y_{1}\right) \gamma\left(y_{3}\right)$, and $h_{2}\left(y_{2}, y_{3}\right)=\beta^{\prime}\left(y_{2}\right) \gamma^{\prime}\left(y_{3}\right)$, for some functions $\alpha, \alpha^{\prime}$ on $Y_{1}, \beta, \beta^{\prime}$ on $Y_{2}$, and $\gamma, \gamma^{\prime}$ on $Y_{3}$, all of them being bounded by 1 in absolute value. Then, for every $n$

$$
\int h_{0} \cdot S_{1}^{n} h_{1} \cdot S_{2}^{n} h_{2} d \nu=\int\left(\alpha \alpha^{\prime}\right)\left(y_{1}\right) \cdot\left(\beta \beta^{\prime}\right)\left(y_{2}\right) \cdot\left(\gamma \gamma^{\prime}\right)\left(R_{3}^{n} y_{3}\right) d \nu\left(y_{1}, y_{2}, y_{3}\right)
$$

By Proposition 4.4, this is equal to

$$
\int \mathbb{E}\left(\alpha \alpha^{\prime} \mid Z\right)\left(z_{1}\right) \mathbb{E}\left(\beta \beta^{\prime} \mid Z\right)\left(z_{2}\right) \mathbb{E}\left(\gamma \gamma^{\prime} \mid Z\right)\left(z_{3}+n \alpha\right) d m\left(z_{1}, z_{2}, z_{3}\right)
$$

Set

$$
U=\left\{t, \int\left|\mathbb{E}\left(\gamma \gamma^{\prime} \mid Z\right)\left(z_{3}+t\right)-\mathbb{E}\left(\gamma \gamma^{\prime} \mid Z\right)\left(z_{3}\right)\right|^{2} d \theta\left(z_{3}\right)<\epsilon\right\}
$$

then $U$ is an open neighborhood of 0 in $Z$ and $\left|J_{0}-J_{n}\right|<\epsilon$ when $n \alpha \in$ $U$.

By linearity, the neighborhood $U$ with the announced property exists when each $h_{i}, i=1,2,3$, is a finite sum of product functions of the type described above. The general case follows by a density argument.

Let $\epsilon>0$ be as in Theorem 4.1, and let $U$ be the neighborhood of 0 in $Z$ defined by Proposition 4.6 with $\epsilon / 2$ substituted for $\epsilon$. Let $\chi$ be a non-negative continuous function on $Z$, such that

$$
\int \chi d \theta=1
$$

and

$$
\chi(z)=0 \text { if } z \notin U .
$$

We keep the notation $\epsilon, U, \chi$ in the sequel.
By Proposition 4.2 and a density argument, we have
Corollary 4.7. The averages of the difference of

$$
\chi(n \alpha)\left(I_{n}-J_{n}\right)
$$

converge to zero.

## Lemma 4.8.

$$
\begin{equation*}
\limsup _{N-M \rightarrow \infty}\left|\frac{1}{N-M} \sum_{n \in[M, N)} \chi(n \alpha)\left(J_{0}-J_{n}\right)\right| \leq \frac{\epsilon}{2} . \tag{27}
\end{equation*}
$$

Proof. If $n \alpha \notin U$, then $\chi(n \alpha)\left(J_{0}-J_{n}\right)=0$. If $n \alpha \in U$, then by Proposition 4.6,

$$
\chi(n \alpha)\left|J_{0}-J_{n}\right|<\frac{\epsilon}{2} \cdot \chi(n \alpha) .
$$

Since $\{n \alpha\}$ is uniformly distributed in $Z$, the averages of $\chi(n \alpha)$ converge to 1 , this gives (27).
4.5. End of the proof. We have the following estimate of $J_{0}$ :

Corollary 4.9 (of Lemma 1.6).

$$
J_{0} \geq\left(\int g d \nu\right)^{4}=\left(\int f d \mu\right)^{4} .
$$

Proof. Since $0 \leq g \leq 1$, we have

$$
J_{0}=\int g_{0} \cdot g_{1} \cdot g_{2} d \nu \geq \int g \cdot g_{0} \cdot g_{1} \cdot g_{2} d \nu
$$

Applying Lemma 1.6 to the probability space $(Y, \nu)$ with $k=3, \mathcal{X}_{1}=$ $Y_{1} \times Y_{2}, \mathcal{X}_{2}=Y_{1} \times Y_{3}$ and $\mathcal{X}_{3}=Y_{2} \times Y_{3}$, the announced conclusion follows.

We can now resume the proof of Theorem 4.1:
Proof of Theorem 4.1. By Corollary 4.9, it suffices to prove that the set $E:=\left\{n \in \mathbb{Z}: I_{n}>J_{0}-\epsilon\right\}$ is syndetic. Suppose by contradiction that $E$ is not syndetic. Then there exists a sequence of intervals $\left[M_{i}, N_{i}\right)$ whose lengths $N_{i}-M_{i}$ tend to infinity and such that $I_{n} \leq J_{0}-\epsilon$ for every $i$ and every $n \in\left[M_{i}, N_{i}\right)$. Therefore

$$
\begin{equation*}
\frac{1}{N_{i}-M_{i}} \sum_{n \in\left[M_{i}, N_{i}\right)} \chi(n \alpha)\left(J_{0}-I_{n}\right) \geq \frac{1}{N_{i}-M_{i}} \sum_{n \in\left[M_{i}, N_{i}\right)} \chi(n \alpha) \cdot \epsilon \tag{28}
\end{equation*}
$$

for every $i$. Taking the lim sup of both sides, since the averages of $\chi(n \alpha)$ converge to 1 , we get that the limsup of the left hand side of (28) is greater than or equal to $\epsilon$.

On the other hand, it follows from Corollary 4.7 and Lemma 4.8 that the lim sup of the left hand side of (28) is less than or equal to $\epsilon / 2$. We have a contradiction.

## 5. Proof of Theorem 1.2

Let $(Y, \nu, S)$ be the two-side $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$-Bernoulli-shift. This means that $\{0,1,2\}$ is endowed with the uniform probability measure $\tau$, and that $Y=\{0,1,2\}^{\mathbb{Z}}$ is endowed with the product measure $\nu=\tau^{\mathbb{Z}}$. The points of $Y$ are written $y=\left(y_{n}: n \in \mathbb{Z}\right)$ where $y_{n} \in\{0,1,2\}$ for every $n$. The shift map $S$ is given by $(S y)_{n}=y_{n+1}$ for every $n$.

Define

$$
\left(X, \mu, T_{1}, T_{2}\right)=(Y \times Y \times Y, \nu \times \nu \times \nu, S \times \operatorname{Id} \times S, \operatorname{Id} \times S \times S)
$$

Since $(Y, \nu, S)$ is weakly mixing, the system $\left(X, \mu, T_{1}, T_{2}\right)$ is ergodic.
For $(i, j, k) \in\{0,1,2\}^{3}$, let $f(i, j, k)$ be defined by the following table:

|  | $\mathrm{i}=0$ | $\mathrm{i}=1$ | $\mathrm{i}=2$ |
| :---: | :---: | :---: | :---: |
| $\mathrm{j}=0$ | 001 | 001 | 111 |
| $\mathrm{j}=1$ | 011 | 001 | 110 |
| $\mathrm{j}=2$ | 110 | 111 | 100 |
| $\mathrm{k}=$ | 012 | 012 | 012 |

Let $F$ be the function on $Y \times Y \times Y$ defined by $F(y, z, w)=f\left(y_{0}, z_{0}, w_{0}\right)$. Then $F$ is an indicator function of some measurable subset $A \subset X$. For every $n \neq 0$, we have

$$
\begin{aligned}
& \mu\left(A \cap T_{1}^{n} A \cap T_{2}^{n} A\right)=\int F \cdot T_{1}^{-n} F \cdot T_{2}^{-n} F d \mu \\
= & \int F(y, z, w) \cdot F\left(S^{-n} y, z, S^{-n} w\right) \cdot F\left(y, S^{-n} z, S^{-n} w\right) d \mu(y, z, w) \\
= & \int f\left(y_{0}, z_{0}, w_{0}\right) \cdot f\left(y_{-n}, z_{0}, w_{-n}\right) \cdot f\left(y_{0}, z_{-n}, w_{-n}\right) d \mu(y, z, w) .
\end{aligned}
$$

Since $n \neq 0$, the variables at the position 0 are independent with those at the position $-n$, we have

$$
\begin{aligned}
& \mu\left(A \cap T_{1}^{n} A \cap T_{2}^{n} A\right) \\
= & \int f\left(y, z, w^{\prime}\right) f\left(y^{\prime}, z, w\right) f\left(y, z^{\prime}, w\right) d \tau(y) d \tau\left(y^{\prime}\right) d \tau(z) d \tau\left(z^{\prime}\right) d \tau(w) d \tau\left(w^{\prime}\right) \\
= & \frac{1}{3^{6}} \sum_{i, j, k, i^{\prime}, j^{\prime}, k^{\prime}} f\left(i, j, k^{\prime}\right) f\left(i, j^{\prime}, k\right) f\left(i, j^{\prime}, k\right) .
\end{aligned}
$$

By computation, for every integer $n \neq 0$, we have

$$
\begin{aligned}
& \mu\left(A \cap T_{1}^{n} A \cap T_{2}^{n} A\right)=\frac{145}{729}<0.96\left(\frac{16}{27}\right)^{3} \\
= & 0.96\left(\frac{1}{3^{3}} \sum_{i, j, k} f(i, j, k)\right)^{3}=0.96(\mu(A))^{3} .
\end{aligned}
$$

Let $0<c \leq 1$. There exists an integer $l>0$ such that $(0.96)^{l}<c$. Let ( $X^{\prime}, \mu^{\prime}, T_{1}^{\prime}, T_{2}^{\prime}$ ) be the $l$-times product system of ( $X, \mu, T_{1}, T_{2}$ ), and $A^{\prime}$ be the $l$-times product of set $A$. Then the system $\left(X^{\prime}, \mu^{\prime}, T_{1}^{\prime}, T_{2}^{\prime}\right)$ is ergodic and the set $A^{\prime}$ satisfies the announced property.

## References

1. T. Austin. On the norm convergence of nonconventional ergodic averages. To appear, Ergodic Theory Dynam. Systems.
2. F. V. Atkinson, G. A. Watterson and P. A. P. Moran. A matrix inequality. Quart. J. Math. Oxford Ser. (2) 11 (1960) 137-140.
3. V. Bergelson and A. Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems. J. Amer. Math. Soc. 9 (1996), no. 3, 725-753.
4. V. Bergelson, B. Host and B. Kra, with an Appendix by I. Ruzsa. Multiple recurrence and nilsequences. Inventiones Math. 160 (2005), 261-303.
5. Q. Chu. Convergence of multiple ergodic averages along cubes for several commuting transformations. Studia Math. 196 (1) (2009), 13-22.
6. Q. Chu, N. Frantzikinakis and B. Host. Commuting averages with polynomial iterates of distinct degrees. Preprint.
7. N. Frantzikinakis. Multiple ergodic averages for three polynomials and applications. Trans. Amer. Math. Soc. 360 (2008), no. 10, 5435-5475.
8. N. Frantzikinakis and B. Kra. Ergodic averages for independent polynomials and applications. J. London Math. Soc. (2) 74 (2006), no. 1, 131-142.
9. H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. J. Anal. Math. 31, 204-256 (1977).
10. H. Furstenberg. Recurrence in ergodic theory and combinatorial number theory. M. B. Porter Lectures. Princeton University Press, Princeton, N.J., (1981).
11. H. Furstenberg and Y. Katznelson. An ergodic Szemerédi theorem for commuting transformations. J. Analyse Math. 34 (1978), 275-291 (1979).
12. B. Host. Ergodic seminorms for commuting transformations and applications. Studia Math. 195 (1) (2009), 31-49.
13. B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. Ann. of Math. 161 (2005), no. 1, 397-488.
14. A.Y. Khintchine. Eine Verschärfung des Poincaréschen "Wiederkehr-Satzes". Compos. Math. 1, (1934), 177-179.
15. A. Sidorenko. A correlation inequality for bipartite graphs. Graphs Combin. 9 (1993), no. 2, 201-204.
16. A. Sidorenko. An analytic approach to extremal problems for graphs and hypergraphs. Extremal problems for finite sets (Visegrád, 1991), 423-455, Bolyai Soc. Math. Stud., 3, János Bolyai Math. Soc., Budapest, 1994.
17. T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. Ergodic Theory Dynam. Systems. 28 (2008), 657-688.
18. H. Towsner. Convergence of diagonal ergodic averages. Ergodic Theory Dynam. Systems 29 (2009), no. 4, 1309-1326.

Université Paris-Est, Laboratoire D'Analyse et de Mathématiques Appliquées, UMR CNRS 8050, 5 bd Descartes, 77454 Marne la Vallée Cedex 2, France

E-mail address: qing.chu@univ-mlv.fr


[^0]:    Date：December 17， 2009.
    2000 Mathematics Subject Classification．37A05，37A30．
    Key words and phrases．Multiple recurrence，commuting transformations，magic system，ergodic seminorms．

