Speedups of compact group extensions

Andrey Babichev and Adam Fieldsteel

Abstract. Let $S_1$ and $S_2$ be ergodic extensions of finite measure preserving transformations $T_1$ and $T_2$, where the extensions are by rotations of a compact group $G$. Then there is an $N$–valued function $k$, measurable with respect to the factor $T_1$, so that $S_k^1$ is isomorphic to $S_2$ by an isomorphism that respects the action of $G$ on fibers.

1. Introduction

Fix a compact group $G$ with Haar measure $\lambda$ and two-sided invariant metric $\rho \leq 1$. Let $T$ be a measure preserving transformation of the Lebesgue probability space $(X, \mathcal{A}, \mu)$ and $\sigma : X \to G$ an $\mathcal{A}$–measurable map. The transformation $S : X \times G \to X \times G$ given by $S(x, g) = (Tx, \sigma(x)g)$ is a measurable map preserving $\mu \times \lambda$. We refer to such an $S$ as a $G$–extension of $T$, or more briefly as a $G$–extension, if $T$ is either understood or need not be specified. The factor $T$ will be referred to as the base factor of $S$ and we will frequently identify the sets in $\mathcal{A}$ with their preimages in $X \times G$ under the projection on the first coordinate. We will use the notation $(S, T, X, \sigma)$ to denote such a $G$–extension, and we will use abbreviations such as $(T, \sigma)$ or $S$ when the other components are understood. We will adopt the notational convention that all $G$–extensions are represented by the letter $S$, or a modified letter $S$, and the associated base factor and function into $G$ will be represented by the letters $T, X, \text{ and } \sigma$, respectively, with the same modifiers. Thus a $G$–extension $S'$ is understood to be associated with the components $(T', X', \mathcal{A}', \mu', \sigma')$. We let $c$ denote the projection $c : (x, g) \mapsto g$.

Each $G$–extension admits a natural free action of $G$ on $X \times G$, which, for each $h \in G$, is given by

$$h(x, g) = (x, gh)$$

and this action commutes with the action of $\mathbb{Z}$ given by (the powers of) $S$. For each $(x, g) \in X \times G$, we refer to the set $G(x, g) = \{(x, gh) \mid h \in G\}$ as the $G$–orbit or the $G$–fiber of $(x, g)$.

Given two $G$ extensions $(S, T, X, \sigma)$ and $(\bar{S}, \bar{T}, \bar{X}, \bar{\sigma})$ we say $S$ is a $G$–factor of $\bar{S}$ if there is a factor map $\Phi$ from $\bar{S}$ to $S$ of the form

$$\Phi(\bar{x}, g) = (\phi(\bar{x}), \bar{\alpha}(\bar{x})g)$$

where $\phi$ is a factor map from $\bar{T}$ to $T$ and $\bar{\alpha} : \bar{X} \to G$ is an $\bar{\mathcal{A}}$–measurable function. If such a $\Phi$ exists for which $\phi$ is an isomorphism from $\bar{T}$ to $T$, we say $S$ is $G$–isomorphic to $\bar{S}$. We note that these relations can be described in terms of
cocycles on equivalence relations. We will not make use of this language, so we omit
the definitions, but we simply state: given a $G$–extension $(S,T,X,\sigma)$, the function
$\sigma$ determines (and is determined by) a $G$–valued cocycle on the orbit relation
of $T$. The condition that two $G$–extensions are $G$–isomorphic says that the base
transformations are isomorphic, and after this identification of the orbit relations
of the base transformations, their associated $G$–cocycles are cohomologous. The
function $\bar{\alpha}$ is the “transfer function” that relates the two cocycles.

By a speedup of a transformation $T : X \to X$ we mean a transformation
$T' : X \to X$ of the form $T'(x) = T^{k(x)}(x)$, for some measurable $k : X \to \mathbb{N}$. Given a $G$–extension $(S,T,X,\sigma)$ we consider speedups of $S$ for which the variable
exponent $k$ is measurable with respect to the base factor, and we refer to such a transformation as a $G$–speedup of $S$. Each $G$–speedup of $S$ determines, and is
determined by, a speedup of the base factor $T$. Thus a $G$–speedup of $S$ can be
understood to be a $G$–extension $S'$ of the form
$$S'(x,g) = (T'x, \sigma'(x)g),$$
where
$$T'(x) = T^{k(x)}(x)$$
for an $\mathcal{A}$–measurable function and $k : X \to \mathbb{N}$, and

(1.1) \[ \sigma'(x) = \sigma^{(k)}(x) = \sigma \left( T^{(k-1)(x)} \right) \ldots \sigma (Tx) \sigma (x). \]

Our goal here is to prove that for all ergodic $G$–extensions $S$ and $\bar{S}$, $S$ can be
obtained as a $G$–speedup of $\bar{S}$. That is, there is a $G$–speedup of $\bar{S}$ that is
$G$–isomorphic to $S$. We note that the restriction of this theorem to the special case
where the group $G$ is trivial is a result obtained by Arnoux, Ornstein and Weiss
[\text{AOW}], and our work here gives a new proof of that result.

The theorem is an analogue of the orbit equivalence result for $G$–extensions obtained in [\text{F}] and independently by other methods in [\text{C}]. The proof will fall into
two main parts. First we will show that, given such $S$ and $\bar{S}$, there is an ergodic
$G$–speedup of $\bar{S}$ that has $S$ as a $G$–factor. We will then improve this result
to obtain an isomorphism. From a broader point of view, the overall argument is
carried out by an argument that is closely related to those of the theory of restricted
orbit equivalence developed by Rudolph and Kammeyer [\text{R}, \text{KR1}, \text{KR2}], and
that is ultimately derived from Ornstein’s proof of the isomorphism theorem for
Bernoulli shifts [\text{O}].

The general idea of the proof is a natural one, which may be obscured by
its implementation. Briefly, to obtain a speedup of a transformation $T$ that is
isomorphic to a transformation $\bar{T}$ we must advance along $\bar{T}$–orbits so that, with
respect to a suitable partition $\bar{P}$, we visit the elements of $P$ in a manner that
imitates the behavior of the orbits of $T$ with respect to a generating partition $P$. The
ergodicity of $\bar{T}$ will make this possible. To obtain a $G$–speedup of a $G$–extension
$\bar{S}$ that is isomorphic to a given $S$, we do the same, with the additional requirement
that we advance along $\bar{S}$–orbits by amounts that are constant on $G$–fibers, in a
manner that imitates the behavior of the orbits of $S$, with respect to both the first
and second coordinates.

A particular technical issue that will concern us here, which was not present
in the earlier work on orbit equivalence [\text{F}], is that of establishing the ergodicity of
our speedups. Ergodicity is preserved under orbit equivalence, but the orbits of a
speedup are suborbits of an ergodic transformation, so special effort will be needed to ensure that the speedups we construct are ergodic. To simplify matters a bit, the main argument will be carried out first in the case of finite partitions and then extended to allow countable partitions.

2. Preliminaries

2.1. Partial transformations. The speedups of the theorem will be obtained as limits of partially defined transformations, which we now introduce.

Definition 1. A partial transformation $T$ on $X$ is an injective, measure-preserving map $T : \text{Dom}(T) \to X$ defined on a measurable subset $\text{Dom}(T)$ of $X$. For such $T$ and for $n \in \mathbb{Z}$ we obtain a partial transformation $T^n$ in a natural way. For each set $C \subset \mathbb{Z}$ and $x \in X$ we let $T^C x = \{T^nx \mid n \in C \text{ and } x \in \text{Dom}(T^n)\}$. In particular, we refer to $T^\mathbb{Z} x$ as the $T$–orbit of $x$. A partial $G$–extension (of a partial transformation $T$) is a map $S : \text{Dom}(T) \times G \to X \times G$ of the form

$$S(x,g) = (Tx, \sigma(x)g)$$

where $\sigma : \text{Dom}(T) \to G$ is $A$–measurable.

Definition 2. A partial speedup of a transformation $T_0 : X \to X$ is a partial transformation $T : \text{Dom}(T) \to X$ that satisfies

$$T(x) = T_0^{k(x)}(x)$$

for all $x \in \text{Dom}(T)$, where $k : \text{Dom}(T) \to \mathbb{N}$ is a measurable function. If $(S_0,T_0,X,\sigma)$ is a $G$–extension then a partial $G$–speedup of $S_0$ is a partial speedup $S$ of $S_0$ where the domain of $S$ is a measurable set of the form $\text{Dom}(S) = X' \times G$, and $S$ has the form

$$S(x,g) = S_0^{k(x)}(x,g)$$

where $k : X' \to \mathbb{N}$ is an $A$–measurable function. Equivalently, we can view $S$ as the partial $G$–extension of the partial speedup $T$ of $T_0$ with domain $X'$, where $T$ is given by the same exponent $k$, and $T$ is extended by the function $\sigma^{(k)}$ as in Definition 3.

Definition 3. If $(S,T,\sigma,X)$ is a $G$–extension and $\alpha : X \to G$ is measurable then we let $(S^\alpha,T,\sigma^\alpha,X)$ denote the $G$–extension given by setting

$$\sigma^\alpha(x) = \alpha(Tx)\sigma(x)\alpha^{-1}(x)$$

We note that $(S^\alpha,T,\sigma^\alpha,X)$ is $G$–isomorphic to $(S,T,\sigma,X)$ via the isomorphism

$$(x,g) \mapsto (x,\alpha(x)g).$$

A similar definition is made, using the same notation, in the case that $(S,T,\sigma,X)$ is a partial $G$–extension, with $\alpha : \text{Dom}(T) \to G$.

2.2. Distributions and Sampling. Let $\mathcal{M}(M)$ denote the space of Borel probability measures on the metric space $(M,\rho)$, where $(M,\rho)$ is taken to be separable, and $\rho$ is bounded by 1. These conditions on $M$ will be understood to be in effect throughout this paper. We make use of the Kantorovich metric on $\mathcal{M}(M)$ (which yields the weak topology on $\mathcal{M}(M)$) defined by setting, for all $\lambda_1, \lambda_2 \in \mathcal{M}(M)$,

$$\|\lambda_1, \lambda_2\|_{\mathcal{M}} = \inf \left\{ \int_{M \times M} \rho(x,y) \, d\nu(x,y) \right\}$$
where the infimum is taken over all probability measures \( \nu \) on \( M \times M \) having marginals \( \lambda_1 \) and \( \lambda_2 \).

Given a (Borel) measurable function \( f \) from a Lebesgue space \((X, \mu)\) to \((M, \rho)\), by the distribution of \( f \), denoted \( \text{dist}_X(f) \) we mean the image of \( \mu \) under \( f \). If \( Y \) is a subset of \( X \) with \( \mu(Y) > 0 \), we obtain \( \text{dist}_Y(f) \) by restricting \( f \) to the normalized measure space \( \left( \frac{Y \cap X}{\mu(Y)} \right) \). If \( A \) is a finite subset of \( X \), we obtain \( \text{dist}_A(f) \) by restricting \( f \) to the space \((A, \nu)\), where \( \nu \) is normalized counting measure on \( A \). When \( M \) is a finite or countable set and no other metric is specified, it will be understood that \( \rho \) is the discrete metric on \( M \).

For \( K \in \mathbb{N} \) we let \([K]\) denote the set \( \{0, 1, \ldots, K - 1\} \). Given \( n \) and \( K \in \mathbb{N} \) and a sequence \( s : [K] \to M \), we obtain a function \( s_n : [K - n + 1] \to (M^n)^{K-n+1} \) by setting, for each \( i \in [K - n + 1] \),

\[
s_n(i) = (s(i), \ldots, s(i + n - 1)).
\]

We refer to \( \text{dist}_{[K-n+1]}(s_n) \) as the \( n \)-distribution of \( s \). The restrictions of \( s \) to intervals of length \( n \) are called \( n \)-blocks in \( s \). More generally, if we specify a set of \( n \)-blocks in \( s \), where \( I \subset [K - n + 1] \) is the set of initial positions of these blocks, then we refer to \( \text{dist}_I(s_n) \) as the \( n \)-distribution of this set of \( n \)-blocks. (We will make use of this especially in the case where the specified \( n \)-blocks (that is, their domains) are pairwise disjoint). The sequences to which we apply this language will often be the values of a function \( f \) along an orbit of a transformation \( T \). In that case the sequence \( \{ f(T^ix) \}_{i=0}^n \) will be referred to as the \( T - f - n \)-name of \( x \). We will also use this language in connection with orbits themselves. In particular, if \( S \) is a speedup of \( S_0 \), we may need to speak about blocks in \( S \)-orbits as well as blocks in \( S_0 \)-orbits, so to distinguish them, we will refer to \( S \)-blocks and \( S_0 \)-blocks.

The Birkhoff ergodic theorem can be formulated as:

**Ergodic Theorem:** Let \( T \) be an ergodic measure preserving transformation of \((X, \mu)\), and \( f : X \to (M, \rho) \) a measurable function. Then for almost every \( x \in X \),

\[
\lim_{n \to \infty} \left\| \text{dist}_{T^n(x)}(f), \text{dist}_X(f) \right\|_\mathcal{M} = 0.
\]

When

\[
\left\| \text{dist}_{T^n(x)}(f), \text{dist}_X(f) \right\|_\mathcal{M} < \zeta
\]

we say that the \( T - f - n \)-name of \( x \) has \( \zeta \)-good distribution. If the function \( f \) has the form

\[
f = \bigvee_{i \in [K]} T^{-i}g
\]

then in the above situation we would say that the \( T - g - n \)-name of \( x \) has \( \zeta \)-good \( k \)-distribution.

The following two lemmas provide key combinatorial devices that will be used in our argument.

**Lemma 1.** Let \((T, X, \mu)\) be an ergodic transformation and \( f : X \to (M, \rho) \) a measurable function. For all \( n \in \mathbb{N} \) and \( \zeta > 0 \) there exists \( L(n, \zeta) \in \mathbb{N} \) so that for all \( L \geq L(n, \zeta) \), \((1 - \zeta)\)-most points have \( L \)-names that can be \((1 - \zeta)\)-covered by a set of disjoint \( n \)-blocks which has \( \zeta \)-good \( n \)-distribution. In addition, these \( n \)-blocks are organized into groups of consecutive \( n \)-blocks where the concatenation
of these groups has $\zeta$–good $n$–distribution. Moreover, the lengths of these groups can be take to exceed any lower bound given in advance.

**Proof.** Given $n$ and $\zeta$, fix $\xi > 0$ and choose $K > \frac{\xi}{2}$ so that for a set $X_1 \subset X$ with $\mu(X_1) > (1 - \xi)$, and for all $x \in X_1$,

$$\left\| \text{dist}_{T^{|K|}x} \left( \bigvee_{i \in [n]} T^{-j}f \right) , \text{dist}_{X} \left( \bigvee_{i \in [n]} T^{-j}f \right) \right\| < \xi$$

Choose finitely many disjoint sets $\{A_i\}$, with $X_2 := \bigcup A_i \subset X_1$, and with $\mu(X_2) > (1 - \xi)$, and so that for each $A_i$ and all $x, y \in A_i$,

$$\max_{0 \leq j \leq K - 1} \{ \rho(T^j x, T^j y) \} < \xi.$$  

(The elements of $X_2$ are “good $K$–points” whose $K$–orbits are “good $K$–blocks”.) Choose $L$ so that most points have an $L$–orbit which is mostly covered by good $K$–blocks, and which can therefore be $(1 - \xi)$–covered by disjoint good $K$–blocks. Moreover, we may arrange that if these disjoint good $K$–blocks are partitioned into “types” according to the $A_i$ that contains their initial element, then the each type repeats at least $\left( \frac{1}{K} \right)$–many times in the given $L$–orbit.

For each such repeated type of good $K$–block, cyclically divide the occurrences of that type of block into consecutive $n$–blocks. That is, divide the $j^{th}$ occurrence of the type into disjoint $n$–blocks starting at position $[j]_n$, where $[i]_n \in \{0, 1, 2, ..., n\}$ and $[j]_n \equiv j \pmod{n}$. If $\xi$ was chosen sufficiently small, the resulting collection of $n$–blocks covers $(1 - \zeta)$ of the $L$–orbit by disjoint $n$–blocks with good $n$–distribution, and these $n$–blocks are organized into consecutive groups nearly $K$ in length. Since the $K$ blocks were good, if these groups of consecutive $n$–blocks are concatenated, the resulting long block has $\zeta$–good $n$–dist. \(\square\)

We note that lemma 1 can immediately be strengthened so that each of the points, whose existence is asserted by the lemma, has an $L$–name with $\zeta$–good $n$–distribution.

**Lemma 2. (constructing a model name)** Let $(T, X)$ be an ergodic transformation and $f : X \to (M, \mu)$ a measurable function. For all $n \in \mathbb{N}$ and $\zeta > 0$, and for all sufficiently large $n_1$, and for arbitrarily large $L'$, there is a sequence $F \in M^{L'}$ such that:

1. The $n_1$–distribution of $F$ is within $\zeta$ of the distribution of $\bigvee_{i \in [n_1]} T^{-i}f$. That is,

$$\left\| \text{dist}_{\left[ L' - n_1 + 1 \right]}(F_{n_1}), \text{dist}_{X} \left( \bigvee_{i \in [n_1]} T^{-i}f \right) \right\| < \zeta$$

2. $F$ is a union of consecutive $n_1$–blocks, and this set of $n_1$–blocks has $n_1$–distribution within $\zeta$ of the distribution of $\bigvee_{i=0}^{n_1-1} T^{-i}f$. That is, if $I = \{i \in [L'] | i \equiv 0 \pmod{n_1}\}$, then

$$\left\| \text{dist}_{I}(F_{n_1}), \text{dist}_{X} \left( \bigvee_{i \in [n_1]} T^{-i}f \right) \right\| < \zeta$$
(3) Each of the disjoint $n_1$ blocks above is at least $(1 - \zeta) - \text{covered by a set of disjoint } n - \text{blocks which has } n - \text{distribution within } \zeta \text{ of the distribution of } \bigvee_{i \in [n]} T^{-i} f.$

**Proof.** Given $(n, \zeta)$, choose $\zeta_1 > 0$ and let $n_1 \geq L(n, \zeta_1)$ (as defined in lemma 1). Choose $L' > L(n_1, \zeta_1)$ so that, in addition, most points have $L'$ names with $\zeta_1$-good $n_1$-distribution. Fix such a point $x \in X$. Cover a $(1 - \zeta_1)$-fraction of its $L'$-orbit by a set of disjoint $n_1$-blocks which has $\zeta_1$-good $n_1$-distribution, and which blocks are organized into groups of consecutive blocks, where each of which group (as a single sequence) has $\zeta_1$-good $n_1$-distribution. So most of these (disjoint) $n_1$-blocks can be $(1 - \zeta_1)$-covered, disjointly, by a set of $n$-blocks with $\zeta$-good $n$-distribution. Throw out the $\zeta_1$-fraction of $n_1$-blocks that can't be so covered, and throw out the $\zeta_1$-fraction of the orbit between the groups of consecutive $n_1$-blocks, and then push these remaining $n_1$-blocks together. If $\zeta_1$ was chosen sufficiently small, this (modified) orbit has the name we want. □

**Lemma 3.** ([convexity lemma]) Let $(V, \| \|)$ be a normed real vector space, and suppose that $v_1, v_2$ and $v_Q \in V$ and $0 < \zeta \leq \varepsilon$. If $v_Q = (1 - \varepsilon) v_1 + \varepsilon v_2$ and $\| v_1 - v_Q \| < \zeta$, then $\| v_2 - v_Q \| < \frac{\zeta}{\varepsilon}$.

**Proof.** We have $v_Q = v_1 + \varepsilon (v_2 - v_1)$, so $v_Q - v_1 = \varepsilon (v_2 - v_1)$, so $\| v_Q - v_1 \| = \varepsilon \| v_2 - v_1 \|$. Similarly $\| v_2 - v_Q \| = (1 - \varepsilon) \| v_2 - v_1 \|$. So

$$\| v_Q - v_2 \| = \frac{(1 - \varepsilon)}{\varepsilon} \| v_Q - v_1 \| < \frac{(1 - \varepsilon)}{\varepsilon} \zeta < \frac{\zeta}{\varepsilon}. \quad \square$$

This lemma will be applied to probability vectors $v_1, v_2$ and $v_Q$ viewed as elements of $\mathbb{R}^l$ with respect to the $l^1$-norm, where $v_Q$ will be the distribution of a $t$-set partition $Q$ on a probability space, and $v_1$ and $v_2$ will be the conditional distributions of $Q$ on subsets of measure $1 - \varepsilon$ and $\varepsilon$, respectively.

**Lemma 4.** ([Sampling lemma]) For all $n \in \mathbb{N}$, $\delta \in (0, 2^{-n})$ and $\zeta > 0$ there exists $K = K(n, \delta, \zeta) \in \mathbb{N}$ so that given any set $E$ with $|E| \leq 2^n$ and any probability measure $\nu$ on $E$ such that for all $e \in E$, $\nu(e) > \delta$, and any set $D$ with $|D| \geq K$, there exists a function $f$ from $D$ onto $E$ so that $\| \text{dist}_D(f) - \nu \|_{\mathcal{M}} < \zeta$.

**Proof.** Fix $n \in \mathbb{N}$, $\delta > 0$ and $\zeta > 0$. Choose $K \in \mathbb{N}$ with $\frac{1}{K} < \min \left\{ \delta, \frac{\zeta}{\varepsilon} \right\}$. Suppose we are given a set $E$ and measure $\nu$ as above, and a set $D$ with $|D| = K' \geq K$. Partition $[0, 1]$ into subintervals whose lengths equal the measures of the atoms of $\nu$. That is, partition $[0, 1]$ by $\{0 = x_0 < x_1 < ... < x_1 = 1\}$ so that $|x_i - x_{i-1}| = \nu(e_i)$, where $e_i$ is the $i^{th}$ element of $E$. Modify this partition by moving each endpoint $x_i$ to the nearest multiple of $\frac{1}{K'}$ below it. This new partition determines a distribution $\nu_1$ that is $\zeta$-close to $\nu$. But there is a function $f : D \rightarrow E$ whose statistical distribution is exactly $\nu_1$. □

**Lemma 5.** ([Exhaustion lemma]) Suppose that $\delta' > 0$, and $\varepsilon > 0$ are given, and suppose $\zeta < \frac{\delta'}{\varepsilon}$. Then for all $K'$ there exists $N' \in \mathbb{N}$ such that if $(Z, \lambda)$ is a discrete probability space with normalized counting measure $\lambda$ such that $|Z| > N'$ and $Q$ is a finite partition of $Z$, each of whose atoms has $\lambda$-measure at least $\delta'$, and if $\{S_i \subset Z\}_{i=1}^{r}$ is a pairwise disjoint (non-empty) sequence of subsets of $Z$ such that for all $i$ and $j$, $|S_i| = K'$, $\text{dist}_{S_i} Q = \text{dist}_{S_j} Q$, $\| \text{dist}_{S_i} Q - \text{dist}_{S_j} Q \|_{\mathcal{M}} < \zeta$
and \( \lambda(\bigcup_{i=1}^{r} S_i) < 1 - \varepsilon \), then there is an additional set \( S_{r+1} \subset Z \setminus \bigcup_{i=1}^{r} S_i \), with \(|S_i| = K'\), on which \( \text{dist}_{S_{r+1}} Q = \text{dist}_{S_i} Q \).

**Proof.** Choose \( N' > \frac{K'}{(\varepsilon')^2} \). Suppose that \((Z, \lambda)\) and \( Q \) and \( \{S_i\}_{i=1}^{r} \) are as in the statement of the lemma. Writing \( S = \bigcup_{i=1}^{r} S_i \), we would have \( \lambda(Z \setminus S) = \varepsilon' > \varepsilon \), and by lemma 3,

\[
\|\text{dist}_{Z \setminus S} Q, \text{dist}_{Z} Q\|_{\mathcal{M}} < \frac{\varepsilon'}{\varepsilon} < \frac{\zeta}{\varepsilon}.
\]

Since \( \frac{\varepsilon'}{\varepsilon} < \frac{\delta'}{\varepsilon} \), each atom of the trace of \( Q \) on \( Z \setminus S \) has conditional measure at least \( \frac{\varepsilon'}{\varepsilon} \) and (unconditional) measure at least \( \left(\frac{\varepsilon'}{\varepsilon} \right) N' > \left(\frac{\varepsilon'}{\varepsilon} \right) N' > K' \), each atom of the trace of \( Q \) on \( Z \setminus S \) has at least \( K' \) elements. Therefore there is an injection \( g : S_1 \rightarrow Z \setminus S \) so that for all \( s \in S_1 \), \( Q(g(s_1)) = Q(s_1) \), and setting \( g(S_1) = S_{r+1} \) completes the proof.

\( \Box \)

**Remark 1.** The above lemma says that if we are planning to take samples from a discrete uniform measure space which is partitioned by \( Q \), where \( Q \) has finitely many atoms and none of very small measure, and if we are planning to do so using samples of a known size \( K' \), then if the samples will have distribution close enough (within \( \zeta \)) to that of \( Q \), and if the discrete space is large enough compared to \( K' \), we will be able to take repeated samples (without replacement) until the space is nearly exhausted (to within preassigned \( \varepsilon \)).

### 2.3. Weak topology.

We collect here some basic facts about the metric \( \| \|_{\mathcal{M}} \) and the topology it generates. As before, \( G \) denotes a compact group with Haar measure \( \lambda \) and two-sided invariant metric \( \rho \).

**Lemma 6.** Let \( \varepsilon > 0 \) and \( n \in \mathbb{N} \), and suppose that the sequence \( \gamma : [n] \rightarrow G \) satisfies

\[
\|\text{dist}_{[n]} \gamma, \lambda\|_{\mathcal{M}} < \varepsilon.
\]

Then for all \( h \in g \),

\[
\|\text{dist}_{[n]} \gamma h, \lambda\|_{\mathcal{M}} < \varepsilon.
\]

**Proof.** The proof is immediate. \( \Box \)

**Lemma 7.** Let \( A \) be an open subset of \( G \). Then for all \( \varepsilon > 0 \) there exists \( \eta > 0 \) such that for all \( n \in \mathbb{N} \), if \( \gamma : [n] \rightarrow G \) is a (finite) sequence in \( G \) such that

\[
\|\text{dist}_{[n]} \gamma, \lambda\|_{\mathcal{M}} < \eta
\]

then

\[
\frac{1}{n} \sum_{i \in [n]} \chi_A(\gamma(i)) \geq \lambda(A) - \varepsilon
\]

**Proof.** Using the fact that \( \| \|_{\mathcal{M}} \) metrizes the weak topology \( [\mathcal{D}] \), the conclusion is a statement of a well-known fact about the weak topology. (see \( [\mathcal{D}] \)) and holds in general for arbitrary probability measures on a separable metric space \( (M, \rho) \) of finite diameter in the place of \( \text{dist}_{[n]} \gamma \) and \( \lambda \) on \( G \). \( \Box \)

Combining the previous two lemmas gives us the following
Lemma 8. Let $A$ be an open subset of $G$. Then for all $\varepsilon > 0$ there exists $\eta > 0$ such that for all $n \in \mathbb{N}$, if $\gamma : [n] \to G$ is a sequence such that

$$\|\text{dist}_{[n]} \gamma, \lambda\|_M < \eta$$

then for all $h \in G$

$$\frac{1}{n} \sum_{i \in [n]} \chi_A(\gamma(i) h) \geq \lambda(A) - \varepsilon$$

Definition 4. Let $\nu$ be a Borel probability measure on a metric space $(M, \rho)$. A set $A \subset M$ is called a continuity set for $\nu$ if $\nu(A^c) = 0$.

We note that for all $x \in M$ and all $\delta > 0$ there exists $\delta' < \delta$ such that the ball $B_{\delta'}(x)$ is a continuity set. This is because at most countably many of the pairwise disjoint circles $\{x' \mid \rho(x, x') = \delta'\}$ can have positive measure. Therefore, if $M$ is compact, then for every $\delta > 0$ there is a finite partition of $M$ into continuity sets of diameter less than $\delta$.

The following lemma follows quickly from lemma 7 (which applies to more general metric spaces, as we’ve indicated).

Lemma 9. Let $\nu$ be a Borel probability measure on a compact metric space $(M, \rho)$. Let $Q = \{Q_1, \ldots, Q_t\}$ be a finite partition of $M$ into continuity sets for $\nu$. Then for all $\zeta > 0$ there exists $\zeta' > 0$ so that if $\gamma : [n] \to M$ is a sequence with

$$\|\text{dist}_{[n]} \gamma, \nu\|_M < \zeta'$$

then

$$\|\text{dist}_{[n]} (Q(\gamma)) \cdot \text{dist}Q\|_M < \zeta$$

We will occasionally need to implement a distribution match in a concrete way. The following lemmas allow us to do this.

Lemma 10. Let $\nu$ be a Borel probability measure on a compact metric space $(M, \rho)$. Then for all $\zeta \in (0, 1)$ there exists $\zeta' > 0$ so that if $\gamma_1, \gamma_2 : [n] \to M$ are sequences such that, for both $i = 1$ and 2,

$$\|\text{dist}_{[n]} \gamma_i, \nu\|_M < \zeta'$$

then there is a bijection $\phi : [n] \to [n]$ such that for $(1 - \zeta) - \text{most } i \in \{1, 2, \ldots, n\}$, we have

$$\rho(\gamma_1(\phi(i)), \gamma_2(i)) < \zeta.$$
there is a map \( \phi : [n_1] \to [n] \) so that for \((1 - \zeta)\)−most \( i \in [n] \), we have
\[
\rho(\gamma_1(\phi(i)), \gamma_2(i)) < \zeta
\]
and for each \( i \in [n] \),
\[
\left| \frac{\phi^{-1}(i)}{n_1} - \frac{1}{n} \right| < \zeta.
\]
(The last condition can be interpreted as saying \( \phi \) nearly preserves the normalized counting measures on \([n_1]\) and \([n]\)).

**Proof.** The proof is similar to the proof of lemma [10] \( \square \)

We will refer to the maps \( \phi \) of lemmas [10] and [11] as \( \zeta \)−distribution matches between the sequences \( \gamma_1 \) and \( \gamma_2 \).

We will also need the following simple observation.

**Lemma 12.** Suppose that \( \gamma : [n] \to G \) is a (finite) sequence in \( G \) such that
\[
\|\text{dist}_{[n]} \gamma, \lambda\|_M < \eta
\]
and \( \alpha : [n] \to G \) is a sequence such that for all \( i \), \( \rho(\alpha(i), id_G) < \eta \). Then
\[
\|\text{dist}_{[n]} \alpha \gamma, \lambda\|_M < 2\eta
\]

**2.4. Rokhlin lemma and ergodicity.** Our argument will depend in an essential way on the Rokhlin lemma. In particular, we will make use of Rokhlin towers in \( G \)−extensions, where the towers are measurable with respect to the base factor.

**Definition 5.** Given a \( G \)−extension \((S, T, X, \sigma)\), a Rokhlin tower measurable with respect to the base factor \((T, A)\) is a pairwise disjoint sequence of sets \( R = \{S^iB\}_{i \in [K]} \) where each \( S^iB \in A \). The set \( B \) is called the base of the tower and \( K \) its height. If \( \mu \times \lambda \left( \bigcup_{i \in [K]} S^iB \right) > 1 - \zeta \), we call \( R \) a \((1 - \zeta)\)−\( K \)−tower. If \( P \) is an \( A \)−measurable partition, then a \( P \)−column of \( R \) is a sequence \( C = \{S^iB'\}_{i \in [K]} \), where \( B' \subset B \) is an atom of the trace of \( \bigvee_{i \in [K]} S^{-i}P \) on \( B \). The sets \( S^iB' \) are referred to as levels of the column \( C \). A sequence of sets of the form \( \{S^iL\}_{i \in [k]} \) where \( L \) is a level of a column \( C \) is called a column-block of \( C \) (of length \( k \)).

The term **column-block** is used to emphasize the distinction between a block consisting of levels of a column and a block consisting of points in an orbit, when both are in play together during our construction below. To be specific, in the arguments to follow we will have occasion to construct Rokhlin towers of the above type with respect to a \( G \)−extension \( \vec{S}_0 = (\vec{T}_0, \vec{X}, \vec{\sigma}) \), but in the presence of a \( G \)−speedup \( \vec{S} = \vec{S}_0^k \) of \( \vec{S}_0 \), where \( k : \vec{X} \to \mathbb{N} \). Columns in these towers will be constructed so that \( k \) is constant on every level, so that we can speak of column blocks that are consecutive images of a level under the speedup \( \vec{S} \) as opposed to \( \vec{S}_0 \). In this case we will speak of \( \vec{S} \)−**column-blocks**, to distinguish them from \( S_0 \)−**column-blocks**.

All the language introduced above concerning Rokhlin towers, blocks and column-blocks will apply in an obvious way to partial transformations. As before, a prefix may be attached whenever we need to distinguish objects associated with a transformation \( S_0 \) from those associated with a speedup \( S \) of \( S_0 \).

The Rokhlin lemma can be formulated as follows.
Lemma 13. Let $T$ be an ergodic measure preserving transformation of $(X, \mu)$, and $f$ a measurable function from $X$ to the metric space $(M, \rho)$. Then for all $K \in \mathbb{N}$ and $\varepsilon > 0$ there is a $(1 - \varepsilon) - K$-tower $R = \{S^i B\}_{i=0}^{K-1}$ such that
\[
\|\text{dist}_B(f), \text{dist}_X(f)\|_M < \varepsilon.
\]

We will need to arrange that the speedups we construct are ergodic. To do this we will use the following criterion for ergodicity. Recall that a transformation $T'$ is said to be in the full group of $T$ if each orbit of $T'$ is contained in an orbit of $T$.

Lemma 14. Fix a sequence $\{C_i\}_{i=1}^{\infty}$ of measurable sets in the probability space $(X, \mathcal{A}, \mu)$ such that the algebra they generate is dense in the measure algebra of $(X, \mathcal{A}, \mu)$. Suppose that $T$ is a transformation of $(X, \mathcal{A}, \mu)$ and for all $i$ and $j$ such that $\mu(C_i) < \mu(C_j)$ and for all $\varepsilon > 0$ there is a transformation $T'$ in the full group of $T$ such that $\mu(C_j \cap T'(C_i)) > (1 - \varepsilon) \mu(C_i)$. Then $T$ is ergodic.

3. Basic iterative procedure

The key argument of the proof of our theorems is contained in the following Distribution Improvement Lemma. This lemma shows that, given a partial $G$-speedup $S$ of a $G$-extension $S_0$, which approximates a $G$-extension $S$, we can make a small modification of $S$ to obtain a partial $G$-speedup that is a much improved approximation of $S$. The rest of this section will be devoted to proving this lemma. The reader familiar with Ornstein’s proof of the isomorphism theorem for Bernoulli shifts will recognize this as the counterpart of the “fundamental lemma” of that argument. The repeated application of this lemma will quickly lead to proofs of the theorems we want, and these will be found in the final section of the paper.

Before formulating the basic lemma, it will be convenient to introduce some new language. First we describe the basic scheme by which orbits of $S$ will be manipulated to obtain a partial speedup. This is a purely combinatorial construction that we describe in terms of sequences of integers.

Recall that for $n \in \mathbb{N}$, $[n]$ denotes $\{0, 1, ..., n - 1\}$. More generally, for $r \in \mathbb{R}^{\geq 0}$ we let $[r] = \mathbb{Z} \cap [0, r)$.

Let $M < M'$ and $w$ be elements of $\mathbb{N}$. Suppose that $u : [w] \to [M' - M + 1]$ such that for all $s \in [w]$, $u(s + 1) - u(s) \geq M$. For all $s \in [w]$ define $W_s : [M] \to [M']$ by $W_s(j) = u(s) + j$. We refer to $\{W_s\}_{s \in [w]}$ as a system of $w$ windows of length $M$ in $[M']$.

Suppose that $p \in [w]$, and suppose that for each $l \in [p]$ and each $j \in \left[\frac{w-l}{p}\right]$ we are given $t^l_j : [p] \to [M]$. Then we obtain $g^l_j : [p] \to [M']$ given by
\[ g^l_j(i) = W_{p+t^l_j(i)}(i). \]

Suppose further that map $(l, j, i) \mapsto g^l_j(i)$ is injective.

Definition 6. We refer to such a family $\Gamma = \{g^l_j\}$ of increasing subsequences of $[M']$ as a cycle (of $p -$ sequences in $[M']$). For each $l$, we refer to $\{g^l_j\}_{j \in \left[\frac{w-l}{p}\right]}$ as the $l$th pass (through $[M']$) of $\Gamma$. For each $l$ and $j$ we refer to the sequence $g^l_j$ as the $j$th stage of the $l$th pass of $\Gamma$. 

We describe this informally: in each stage of the cycle $\Gamma$ the function $g^l_j$ selects one point from each of a sequence of $p$ successive windows in $[M']$. The $l^{th}$ pass of $\Gamma$ is a sequence of stages, whose 0$^{th}$ stage begins in window $\tilde{W}_l$, whose every stage begins at the window immediately after the last window of the previous stage, and where as many stages are completed as the sequence of windows can accommodate. When this scheme is implemented below, the sequence $[M']$ will correspond to an orbit block of a transformation, and each $g^l_j$ will identify an orbit of a partial speedup of that transformation.

We note that, for all $s \in [w]$, \[
\left\{(l,j,i) \mid l \in [p], j \in \left[\frac{w-l}{p}\right], i \in [p], \text{ and } jp + l + i = s\right\} \leq p
\]
and more importantly, if $|s| \geq p - 1$ and $|w - s| \geq p - 1$, then the above cardinality equals $p$, and \[
\{i \in [p] \mid (3l,j) \text{ such that } jp + l + i = s\} = [p]
\]
Indeed, if $|s| \geq p - 1$ and $|w - s| \geq p - 1$, then on pass $l$, $g^l_j(i)$ lies in the range of $\tilde{W}_s$ for exactly one value $i$, and in the subsequent pass $l + 1$, the corresponding value of $i$ is one less (mod $p$).

Next we describe the special form that each of the partial speedups that we construct will have. To describe this form, suppose that $(\tilde{S}, \tilde{T}, \tilde{\sigma}, \tilde{X})$ is a partial $G$-speedup of a $G$-extension $(\tilde{S}_0, \tilde{T}_0, \tilde{\sigma}_0, \tilde{X})$ on a space $(\tilde{X} \times G)$. Let $k : Dom(\tilde{T}) \to \mathbb{N}$ denote the measurable function such that for every $\tilde{x} \in Dom(\tilde{T})$, $\tilde{S}(\tilde{x}, g) = \tilde{S}_0^{k(\tilde{x})}(\tilde{x}, g)$ and $\tilde{\sigma} = \tilde{\sigma}_0^{(k(\tilde{x}))}$ the associated "skewing" function as in $\text{[11]}$.

**Definition 7.** If $\tilde{P}$ is a measurable partition of $\tilde{X}$, we will say that the pair $(\tilde{S}, \tilde{P})$ is a regular partial $G$-speedup of $\tilde{S}_0$ if the following conditions are met:

1. There is a set $\tilde{B}$, measurable with respect to $\tilde{X}$, such that for some $L \in \mathbb{N}$, the sets $\{\tilde{S}^i\tilde{B}\}_{i \in [L]}$ are disjoint, and the domain of $\tilde{S}$ is precisely $\bigcup_{i \in [L-1]} \tilde{S}^i\tilde{B}$. (We refer to $\tilde{R} = \bigcup_{i \in [L]} \tilde{S}^i\tilde{B}$ as the speedup tower for $\tilde{S}$ and to $\tilde{B}$ as its base).

2. $k$ is bounded

3. For all $(\tilde{x}, g), (\tilde{x}', g) \in \tilde{B}$,
\[
\bigvee_{i \in [L]} \tilde{S}^{-i}(\tilde{P} \vee \tilde{c})(\tilde{x}, g) = \bigvee_{i \in [L]} \tilde{S}^{-i}(\tilde{P} \vee \tilde{c})(\tilde{x}', g).
\]
(Recall that $\tilde{c}$ denotes the projection on the $G$-coordinate).

If, in addition, for some $n \in \mathbb{N}$ and $\delta > 0$ we have the further properties that

4. $L$ is a multiple of $n$, and for each $(\tilde{x}, g) \in \tilde{B}$,
\[
\left\| dist_{\tilde{S}_0}(\tilde{x}, g) \bigvee_{i \in [n]} \tilde{S}^{-i}(\tilde{P} \vee \tilde{c}), dist_{Dom(\tilde{S}_0)}(\tilde{x}) \bigvee_{i \in [n]} \tilde{S}^{-i}(\tilde{P} \vee \tilde{c}) \right\|_M < \delta.
\]

That is, when the $\tilde{S} - \tilde{L}$-orbit of $(\tilde{x}, g)$ is divided into disjoint, consecutive $n$-blocks, those $n$-blocks have a distribution of names that is $\delta$-close to the full $n$-distribution of the speedup $\tilde{S}$, and
(5) $\bar{\mu} \times \lambda \left( \text{Dom} \left( \bar{S} \right) \right) > 1 - \delta$.
then we will say that the pair $(\bar{S}, \bar{P})$ is $(n, \delta)$–regular.

We note that if $\mu \times \lambda \left( \text{Dom} \left( \bar{S} \right) \right) > 1 - \delta$, then we must have $L > \frac{1 - \delta}{\delta}$. (If $l$ is the measure of a single level of the speedup tower, then $l < \delta$, but $(L - 1) l > 1 - \delta$ so $L > L - 1 > \frac{1 - \delta}{\delta}$.)

DEFINITION 8. If $(\bar{S}, \bar{P})$ is an $(n, \delta)$–regular partial $G$–speedup of $\bar{S}_0$ as above, we refer to the set

$$\Lambda_n \left( \bar{S} \right) = \bigcup_{i \in [\bar{n}]} \bar{S}^i \left( \bar{B} \right)$$

as the $n$–ladder of $\bar{S}$. We refer to a block of the form $\bar{S}^i \left( \bar{x}, g \right)$, where $(\bar{x}, g) \in \Lambda_n \left( \bar{S} \right)$ as a ladder block of $\bar{S}$. Suppose that $(\bar{x}', g')$ is a point in the ladder block $\bar{S}^i \left( \bar{x}, g \right)$, and $\bar{S}$ is another partial transformation on $X \times G$. We say that the ladder block of $(\bar{x}', g')$ is broken by $\bar{S}$ if for some $i \in [0, n - 2]$, $\bar{S} \left( \bar{S}^i \left( \bar{x}, g \right) \right) \neq \bar{S} \left( \bar{S}^i \left( \bar{x}, g \right) \right)$.

LEMMA 15. (Distribution Improvement Lemma) For all $\varepsilon > 0$ and for all open $A_2 \subseteq G$, there exist $\delta > 0$ and $n \in \mathbb{N}$ such that, if $(S, T, \sigma, X)$ and $(\bar{S}_0, \bar{T}_0, \bar{\sigma}_0, \bar{X})$ are ergodic $G$–extensions on $(X \times G)$ and $(\bar{X} \times G)$ respectively, and $(\bar{S}, \bar{T}, \bar{\sigma}, \bar{X})$ is a partial $G$–speedup of $\bar{S}_0$ and $P$ and $\bar{P}$ are finite partitions of $X$ and $\bar{X}$ such that $(\bar{S}, \bar{P})$ is $(n, \delta)$–regular, such that

$$\left\| \text{dist}_{X \times G} \bigcup_{i \in [n]} S^{-i} \left( P \lor c \right), \text{dist}_{\text{Dom}(\bar{S}^n)} \bigcup_{i \in [\bar{n}]} \bar{S}^{-i} \left( \bar{P} \lor \bar{c} \right) \right\|_\mathcal{M} < \delta$$

then for all $\delta_1 > 0$ and all sufficiently large $n_1$, and all $A_1 \subseteq A$, there is a partial $G$–speedup $\bar{S}_1$ of $\bar{S}_0$ and a partition $\bar{P}_1$ and a measurable function $\bar{\alpha} : \bar{X} \rightarrow G$ such that

(3.2) $(\bar{S}_1, \bar{P}_1)$ is $(n_1, \delta_1)$–regular,

(3.3) $|\bar{P} - \bar{P}_1| < \varepsilon$,

(3.4) $\int \rho \left( \bar{\alpha} \left( \bar{x} \right), \text{id}_G \right) d\bar{\mu} \left( \bar{x} \right) < \varepsilon$,

if $D$ denotes the set of points in the speedup tower of $\bar{S}$ such whose ladder block is broken by $\bar{S}_1$, then

(3.5) $\bar{\mu} \times \lambda \left( D \right) < \delta_1$,

(3.6) $$\left\| \text{dist}_{X \times G} \bigcup_{i \in [n_1]} S^{-i} \left( P \lor c \right), \text{dist}_{\text{Dom}(\bar{S}^{n_1}_1)} \bigcup_{i \in [\bar{n}_1]} \bar{S}^{-i}_1 \left( \bar{P}_1 \lor \bar{\alpha} \bar{c} \right) \right\|_\mathcal{M} < \delta_1,$$

and setting $A = A_1 \times A_2$, the set of $y \in \Lambda_{n_1} \left( \bar{S}_1 \right)$ such that

(3.7) $$\frac{1}{n_1} \sum_{i \in [n_1]} 1_A \left( \bar{S}^i_1 \left( y \right) \right) > \left( \bar{\mu} \times \lambda \right) \left( A \right) - \varepsilon$$

has measure greater than $(1 - \varepsilon) \left( \bar{\mu} \times \lambda \right) \left( \Lambda_{n_1} \left( \bar{S}_1 \right) \right)$. 

PROOF. Fix \( \varepsilon > 0 \) and an open set \( A_2 \subset G \). Choose \( \varepsilon' \) as in lemma 8 with respect to \( \frac{1}{100} \) and \( A_2 \). (We may assume that \( \varepsilon' < \min\{1, \varepsilon\} \).) Choose \( n \) and \( \delta \) so that \( n > \frac{1}{100} \) and \( 2^{-(n+1)} > \frac{\varepsilon'}{100} \) and \( \delta < \frac{(\varepsilon')^4}{2^{100}} \). The number \( \delta \) is chosen in part so that the domain of \( S^n \) has measure greater than \( 1 - \frac{\delta}{100} \) and also so that the union of the atoms of measure less than \( \delta \) in any distribution with \( \leq 2^n \) atoms has measure less than \( \frac{(\varepsilon')^4}{100} \). Additional features of the dependence of \( \delta \) on \( \varepsilon' \) and \( n \) will be given below. Suppose the above hypotheses are met concerning the given \( G \)-extensions and the speedup \( \bar{S} \). Let \( \bar{R} \) denote the speedup tower for \( \bar{S} \). We may assume, without loss of generality, that \( (\bar{\mu} \times \lambda)(\bar{R}) < 1 - \frac{\delta}{2} \).

Fix a \( t \)-element partition \( Q \) of \( (P \times G)^{[n]} \) whose atoms are continuity sets of diameter less than \( \frac{n}{100} \) (in the “max” metric \( \rho' \) using the discrete metric on \( P \)) for the measure \( \text{dist}_{X \times G} \bigvee_{i \in [n]} S^{-i} (P \lor c) \). Thus, for all \( q \in Q \), and all \( x, y \in q \) the \( P - n \) names of \( x \) and \( y \) are equal, and the \( c - n \) names are uniformly close.

Let \( n_1 \in \mathbb{N} \) and \( \delta_1 > 0 \) be given. We may assume that \( \delta_1 < \delta \), and we will have occasion to replace \( \delta_1 \) by an even smaller number during the argument. Fix a measurable set \( A_1 \subset \bar{X} \), and let \( A = A_1 \times A_2 \).

Fix \( \zeta > 0 \), whose size will be determined by what follows. Choose \( K \in \mathbb{N} \) by lemma 8 with respect to \( n, \delta \) and \( \zeta \).

Using lemma 2 we fix a model name \( F \) for the “target” process \( (S, P \lor c) \), so that (replacing \( n_1 \) by a larger number if necessary, which we still call \( n_1 \)) \( n_1 > n/\zeta \), \( |F| > n_1/\zeta \) and

(a) \[
\left\| \text{dist}_{|F| - n_1 + 1} F_{n_1}, \text{dist}_{X \times G} \bigvee_{i \in [n_1]} S^{-i} (P \lor c) \right\|_{M} < \frac{\delta_1}{100}
\]

(b) \[
\left\| \text{dist}_{\left[\frac{|F|}{n_1}\right]} n_1 F_{n_1}, \text{dist}_{X \times G} \bigvee_{i \in [n_1]} S^{-i} (P \lor c) \right\|_{M} < \frac{\delta_1}{100}
\]

(c) For every \( k \in \left[\frac{|F|}{n_1}\right] \), \( H_k := F_{n_1} (kn_1) \) is \( (1 - \zeta) \)-covered by disjoint \( n \)-blocks, such that if \( \mathcal{J}_k \) is the set of their initial positions,

\[
\left\| \text{dist}_{\mathcal{J}_k} F_{n_1}, \text{dist}_{X \times G} \bigvee_{i \in [n]} S^{-i} (P \lor c) \right\|_{M} < \tilde{\zeta},
\]

where \( \tilde{\zeta} \) is determined by lemma 8 with respect to \( \zeta \).

(d) For every \( k \), and for every \( q \in Q, \{j \in \mathcal{J}_k \mid Q(F_n(j)) = q\} \geq K \).

We may assume further that for all \( k \) and \( k' \), \( |\mathcal{J}_k| = |\mathcal{J}_{k'}| \). For each \( k > 0 \), we fix a bijection \( \psi_k : \mathcal{J}_k \to \mathcal{J}_0 \) such that \( \|\text{dist}_{\mathcal{J}_k} F_{n_1} \text{dist}_{\mathcal{J}_0} F_{n_1} \psi_k\|_{M} < \zeta \). (To be precise, we must have originally chosen \( \tilde{\zeta} \) as in lemma 8, so that such \( \zeta \)-distribution matches are available here.) We let \( \psi_0 : \mathcal{J}_0 \to \mathcal{J}_0 \) be the identity.

For each \( k \), and each \( J \in \mathcal{J}_k \), we let \( b(J) \) denote the \( n \)-block with initial position \( J \), and refer to \( b(J) \) as a real \( n \)-block of \( H_k \). Each component of \( H_k \setminus \bigcup_{J \in \mathcal{J}_k} b(J) \) is further partitioned into blocks of length no greater than \( n \), using as many blocks of length \( n \) as possible. We refer to these blocks as pseudo \( n \)-blocks. We let \( \mathcal{J}'_k = \{J_{k,i}\}_i \) denote the set of all initial positions of \( n \)-blocks...
(real and pseudo) in $H_k$, listed in order. We also write \( \{J_{k,i_m}\} \) (respectively \( \{J_{k,j_m}\} \)) for the initial positions of the real (respectively pseudo) $n-$blocks in $H_k$, listed in order.

We note that as a consequence of (c.), for every $k$ and $k'$, the distributions of $Q$ on the disjoint $n-$blocks of $H_k$ and $H_{k'}$ are $2\zeta-$close to each other.

The need to replace $n_1$ with a larger number is the reason that the statement of the lemma says “for all sufficiently large $n_1$”.

To support our construction, we use a pair of Rokhlin towers. First, we construct a Rokhlin tower $R$ for $\tilde{S}_0 = \tilde{S}_0^\mu$, with base $B \in \mathcal{A}$ and height $M$ so that $(\tilde{\mu} \times \lambda) (R) > 1 - \zeta$, and for all $y \in B$, $\tilde{S}_0^M (y)$ admits a disjoint collection of ladder blocks of $\tilde{S}$, such that, if $\mathcal{V}_y$ denotes the set of initial elements of these $\tilde{S} - n-$blocks,

\[
|n|_{\mathcal{V}_y} - (\tilde{\mu} \times \lambda) (\tilde{R}) | < \zeta
\]

\[
\left| \text{dist}_{\mathcal{V}_y} \bigvee_{i \in [n]} \tilde{S}^{-i} (\tilde{P} \vee \tilde{c}) \right| < \delta
\]

(3.8)

and for all $y' \in B$

\[
\left| \text{dist}_{\mathcal{V}_y} \bigvee_{i \in [n]} \tilde{S}^{-i} (\tilde{P} \vee \tilde{c} \vee 1_{A_k}) \right| < \zeta
\]

(3.10)

and for all $y, y' \in B$

\[
|\mathcal{V}_y| = |\mathcal{V}_{y'}|
\]

(3.11)

To construct $R$ let $\tilde{B}$ denote the base of $\tilde{R}$ and $\tilde{L}$ its height, and let $\tilde{k}$ denote the variable exponent that gives $\tilde{S} = \tilde{S}_0^\mu$. For each $M \in \mathbb{N}$ and $y \in \bar{X} \times G$ we let $\mathcal{R}_y = \{ y' \in \tilde{S}_0^M (y) : \tilde{S}^\tilde{k} (y') \subset \tilde{S}_0^M (y) \}$ and $\mathcal{Z}_y = \bar{B} \cap \mathcal{R}_y$. Consider the set $\bar{Y} \subset \bar{X} \times G$ consisting of those $y$ such that

\[
|\mathcal{R}_y| - (\tilde{\mu} \times \lambda) (\tilde{R}) | < \zeta
\]

and

\[
\left| \text{dist}_{\mathcal{Z}_y} \bigvee_{i \in [\tilde{L}]} \tilde{S}^{-i} (\tilde{P} \vee \tilde{c} \vee 1_{A_k}) \right| < \zeta
\]

(3.11)

By the ergodic theorem, and using the fact that $\tilde{k}$ is bounded, we know that if $M$ is sufficiently large, then $(\tilde{\mu} \times \lambda) (\bar{Y}) > 1 - \frac{\zeta}{100}$.

Suppose $y \in \bar{Y}$. Then for each $y' \in \mathcal{Z}_y$ the set $\tilde{S}^{[\tilde{L}]} (y')$ is divided into ladder blocks (of length $n$), and if $\mathcal{V}_y$ denotes the initial elements of all these blocks, conditions (3.8), (3.9), and (3.10) are satisfied.

The set $\bar{Y}$ is $A_0-$measurable, so lemma [13] applied to $T_0$ gives a Rokhlin tower $R$ for $\tilde{S}_0$ with base $B \subset \bar{Y}$ that satisfies the desired conditions. By deleting some members of the sets $\mathcal{V}_y$, we can also arrange that (3.11) holds.
In the construction of the tower $R$ we also choose $M$ so that any distribution match as in lemma 11 between $\text{dist}_y \left( \bigvee_{i \in [n]} S^{-i} \left( \tilde{P} \vee \tilde{\epsilon} \vee 1_{A_1} \right) \right)$ for a point $y \in B$ and $\text{dist}_x F_n$ will be at least $\frac{1}{q}$ to-one. We will refer to the $\tilde{S} - n$-blocks with initial points in the sets $\mathcal{V}_y$ as useful blocks.

Second, we construct a much longer Rokhlin tower $R'$ with height $M'$ and base $B' \subseteq \tilde{A}_0$ such that $(\bar{\mu} \times \lambda) (R') > 1 - \zeta$ and so that for each $y \in B'$, $\tilde{S}_0^{[M']} y$ is $(1 - \zeta)$-covered by (necessarily disjoint) sets of the form $\tilde{S}_0^{[M]} y'$ where $y' \in B$. The height $M'$ will be chosen subject to some additional requirements, which will be described below. We divide $R'$ into columns whose levels are pure with respect to $\tilde{P}, 1_{A_1}$, the variable exponent $\tilde{k}$, the levels of the speedup tower $\tilde{R}$ for $\tilde{S}$, and the levels of $R$. We further refine the columns of $R'$ so that in each level $L$ of a column, the values of the skewing function $\tilde{\sigma}_0$ are within $\zeta'$ of being constant, where $\zeta'$ is so small that if $O_{\zeta'}(\text{id}_G)$ denotes the $\zeta'$-ball in $G$ centered at $\text{id}_G$, then

$$
(O_{\zeta'}(\text{id}_G))^{M'} \subseteq O_{\zeta'}(\text{id}_G).
$$

Fix a column $C'$ in $R'$ and fix $y' \in B' \cap C'$. If $j \in [M' - M]$ and $\tilde{S}_0^j y' \in B$ then we refer to $\tilde{S}_0^{[M]} \left( \tilde{S}_0^j y' \right)$ as a window (for $y'$). (This language anticipates the construction of cycles below). We let $\{ W_i \}_{i=0}^{w-1}$ be the set of windows for $y'$, indexed in the order imposed by $\tilde{S}_0$. We recall that each such window is covered, up to a fraction $(\bar{\mu} \times \lambda) (\tilde{R}) \pm \zeta$, by the useful blocks associated with $\tilde{S}_0^j (y)$. We note that the useful blocks in each $W_i$ occupy no more than a $(1 - \frac{\alpha}{q} + \zeta')$ -fraction of $W_i$. We let $\mathcal{V}_i$ denote the set of initial points of these useful blocks.

For each $i > 0$ we fix a bijection $\phi_i : \mathcal{V}_i \to \mathcal{V}_0$ that is a $\zeta'$-distribution match between $\text{dist}_y \left( \bigvee_{i \in [n]} S^{-i} \left( \tilde{P} \vee \tilde{\epsilon} \vee 1_{A_1} \right) \right)$ and $\text{dist}_{\tilde{y}_i} \left( \bigvee_{i \in [n]} S^{-i} \left( \tilde{P} \vee \tilde{\epsilon} \vee 1_{A_1} \right) \right)$. (As before, we must have originally chosen a number $\zeta'$ as in lemma 10 instead of $\zeta$, so that such $\zeta'$-distribution matches are available here).

Let $\theta : \mathcal{V}_0 \to J_0$ be a $\delta$-distribution match between $\text{dist}_y \left( \bigvee_{i \in [n]} S^{-i} \left( \tilde{P} \vee \tilde{\epsilon} \right) \right)$ and $\text{dist}_x F_n$. (Again, we must have chosen a number $\delta = \tilde{\delta} (\delta, n)$ as in lemma 11 instead of $\delta$ in condition 3.1 so that such a $\delta$-distribution match is available here).

The desired partial speedup $\tilde{S}_1 = \tilde{S}_0^{k_1}$ of $\tilde{S}_0$ will be defined first on $\tilde{S}_0^{[M']} (y')$ by concatenating blocks taken from successive windows in $\tilde{S}_0^{[M']} (y')$. These blocks will be selected so that the orbit they form will be well matched to the model name $F$. The definition of $\tilde{S}_1$ will then be extended to the rest of $C'$ by making $k_1$ constant on the levels of $C'$. All other columns will be treated in a similar way. The details will be presented in a sequence of steps.

Step 1. Let $Q_0$ denote the partition $Q \circ \theta$ on $\mathcal{V}_0$. Let $\mathcal{R}_0$ denote the partition of $\mathcal{V}_0$ by $\bigvee_{i \in [n]} S^{-i} (1_{A_1})$. We make a modification $\mathcal{R}_0$ of $\mathcal{R}_0$:

Let $\nu$ denote the normalized counting measure on $\mathcal{V}_0$. For each atom $q$ of $Q_0$, let $\nu_q$ denote $\nu$ conditioned on $q$, and let $\mathcal{R}_0^q$ denote the restriction of $\mathcal{R}_0$ to $q$. We will construct a new partition $\mathcal{R}_0^q$ on $q$, such that $\left\| \text{dist}_q \mathcal{R}_0^q, \text{dist}_q \mathcal{R}_0^q \right\|_M < \frac{\alpha}{q}$, and such that for every atom $\tilde{r}$ of $\mathcal{R}_0^q$, $\nu_q (\tilde{r}) > \delta$.

Let $U_q = \bigcup \{ a \in \mathcal{R}_0^q : \nu_q (a) < \delta \}$. By the choice of $\delta$, $\nu_q (U) < \frac{\delta}{100}$ if $\nu_q (U_q) \geq \delta$. We regard $U_q$ as a single atom of $\mathcal{R}_0^q$ and let $\mathcal{R}_0^q$ coincide with $\mathcal{R}_0^q$ on the rest of its atoms. If $\nu_q (U_q) < \delta$ we choose an atom $r$ of $\mathcal{R}_0^q$ such that $\nu_q (r) > 2^{-n}$ and
a subset \( r' \subseteq r \) with \( \nu_q(r') \in (2^{-(n+2)}, 2^{-(n+1)}) \). We regard \( r' \cup U_q \) and \( r \setminus r' \) as single atoms of \( \mathcal{R}_0^q \), and we let \( \mathcal{R}_0^q \) coincide with \( \mathcal{R}_0^q \) on the rest of its atoms. (The \( \nu_q \)-measure of a singleton is small enough to guarantee the existence of \( r' \)).

The partition \( \mathcal{R}_0^q \) has the desired properties. In particular, \( \mathcal{R}_0^q \) coincides with \( \mathcal{R}_0^q \) except on the atom \( r_{U_q} \) of \( \mathcal{R}_0^q \) that contains \( U_q \).

We let \( \mathcal{R}_0 \) denote the partition of \( \mathcal{V}_0 \) whose restriction to each atom \( q \) is \( \mathcal{R}_0^q \). Thus \( \mathcal{R}_0 \) coincides with \( \mathcal{R}_0 \) except on the union of the sets \( r_{U_q} \). We will refer to the atoms \( r_{U_q} \) as "miscellaneous" atoms.

Step 2. To each real \( n \)-block of \( F \) we assign an atom of \( Q \cap \mathcal{R}_0 \): For each \( q \in Q \), let \( \mathcal{J}_{0,q} = \{ j \in \mathcal{J}_0 \mid Q(F_n(j)) = q \} \). Let \( f_{q,0} : \mathcal{J}_{0,q} \to \mathcal{R}_0^q \) be a function with statistical distribution within \( Q \) of \( \text{dist}_q \mathcal{R}_0^q \). Lemma 4 guarantees that such a function exists. Let \( f_0 : \mathcal{J}_0 \to Q \cap \mathcal{R}_0 \) be the common extension of the \( f_{q,0} \). That is, \( f_0 = \cup_q f_{q,0} \).

For each \( k > 0 \), we set \( f_k = f_0 \circ \psi_k : \mathcal{J}_k \to Q \cap \mathcal{R}_0 \), and we set \( f = \cup_k f_k \).

Step 3. We prepare samples from each \( \mathcal{V}_i \) : Let \( \{ \tau_{0,t} : \cup_k J_k \to \mathcal{V}_0 \} \) be injections with disjoint ranges so that for all \( t \) and all \( i \in \cup_k J_k \), \( Q \cap \mathcal{R}_0^q (\tau_{0,t}(i)) = f(i) \). To see that a large collection of such "samples" \( \tau_{0,t} \) is available, let \( \delta' = \frac{1}{2} \min_{q \in Q} \{ \mu \times \lambda(q) \} \) and \( \delta'' = \delta' \delta \). If \( \zeta \) was chosen so that \( \zeta < \min \left\{ \frac{\delta''}{200}, \frac{\delta'' \delta'}{200} \right\} \), then by property (c) in our choice of \( F \), it will be the case that \( \delta'' < \min_{q \in Q} \nu(q) \), and so \( \delta < \min_{q \in Q \cap \mathcal{R}_0} \nu(q) \). Let \( N' \) be the number given by lemma 5 with respect to \( \delta', \frac{\delta'}{200} \), \( \zeta \) and \( K' = |\cup_k J_k| \). Then the tower \( R \) could have been chosen so that the number of useful blocks in each window exceeds \( N' \). Applying lemma 5 we obtain a set of samples \( \{ \tau_{0,t} \}_{t} \) such that \( \cup_t \tau_{0,t}(\cup_k J_k) \) covers all but a \( \frac{\delta}{100} \)-fraction of \( \mathcal{V}_0 \).

For each \( s > 0 \), we set \( \tau_{s,t} = \phi_s^{-1}\tau_{0,t} : \cup_k J_k \to \mathcal{V}_s \), which provides corresponding samples of \( \mathcal{V}_s \).

Step 4. We construct cycles in \( \mathcal{V}' \) : Identifying \( S_0^{[M']} \) with \( [M'] \) via \( e : S_0^{[M']} \to i \), we have a system \( \{ \tilde{W}_s \} \) of \( w \) windows of length \( M \) in \( [M'] \). Let \( p = |J'| \) be the number of \( n \)-blocks in \( F \). We also identify \( J' \) with \( [p] \) (without changing notation), and for each \( t \), we construct a cycle \( \Gamma_t \) of \( p \)-sequences in \( [M'] \) as follows. Let \( \tau_{s,t}(i) \) denote the height of the point \( \tau_{s,t}(i) \) above the base of \( W_s \). We (partially) define \( \Gamma_t = \{ g_{j,t}^{(l,t)} \}_{l,j} \) by setting,

\[
g_{j,t}^{(l,t)}(i) = \tilde{W}_{jp+l+1} \left( \tau_{jp+l+1,t}^{(l,t)}(i) \right).
\]

For the moment, we leave the functions \( g_{j,t}^{(l,t)} \) of \( \Gamma_t \) undefined on \( J' \cup_k J_k \). We will refer to these 4-tuples \( (t,l,j,i) \) as the real 4-tuples (associated with \( C' \)). We have that \( \{ g_{j,t}^{(l,t)}(i) \mid (t,l,j,i) \text{ is a real } 4 \text{-tuple} \} \) covers all but a \( \frac{\delta}{100} \)-fraction of \( e(\cup \mathcal{V}_s) \).

Step 5. We extend the domains of the functions that make up the cycles \( \Gamma_t \) to all of \( [p] \) : For each window \( W_s \) and each \( z \in \mathcal{V}_s \), let \( b(z) \) denote the useful block beginning at \( z \). The set \( U_s = W_s \setminus \bigcup_{l,j,i} b(\epsilon^{-1}g_{j,t}^{(l,t)}(i)) \) covers at least a \( \left( \frac{\delta}{2} - \epsilon \right) \)-fraction of \( W_s \). For each 4-tuple \( (t,l,j,i) \) where \( (t,l,j) \) is the initial triple of a real 4-tuple and \( i \) is initial position of a pseudo \( n \)-block of length \( n \), we choose a subset \( \gamma((t,l,j,i)) \subset U_s \) of size \( n \). The sets \( \gamma((t,l,j,i)) \) are chosen
to be pairwise disjoint. Since the number of pseudo \( n \)-blocks in \( F \) is at most a \( \zeta \)-fraction of the number of real \( n \)-blocks, and \( \zeta \) is much less than \( \delta \), such disjoint sets are available. We define \( g_j^{l,t} (i) \) to be the first element of \( e (\gamma ((t, l, j, i))) \).

Step 6. We begin to define \( S_1 \): For each stage \( g_j^{l,t} \) in one of the constructed cycles let \( \tilde{g}_j^{l,t} = e^{-1} g_j^{l,t} \). Each \( i' \in J' \) is the initial position of an \( n \) block (real or pseudo) of length \( n \), and \( \tilde{g}_j^{l,t} (i') \) is the initial point of a block of the same length \( \tilde{n} \) in a single window of \( S_0^{[M']} \) \( (y') \). That block is either the useful block beginning at \( \tilde{g}_j^{l,t} (i') \) or a block of the form \( \gamma ((t, l, j, i)) \) as in step 5. The blocks with initial points \( \tilde{g}_j^{l,t} (i') \) lie in successive windows, and we extend the map \( \tilde{g}_j^{l,t} \) to a map \( \tilde{g}_j^{l,t} : |F| \to \tilde{S}_0^{[M']} (y') \) by concatenating these blocks. It follows that \( e\tilde{g}_j^{l,t} : |F| \to [M'] \) is increasing. We regard the image of \( \tilde{g}_j^{l,t} \) as an orbit of the new partial speedup \( \tilde{S}_1 \). Writing \( \bar{S}_1 = \bar{S}_0^{k_1} \) along this orbit, we extend the definition of \( \bar{S}_1 \) to the union of the levels of \( C' \) that contain this orbit, by requiring that \( \bar{k}_1 \) be constant on each of these levels. We refer to this union of levels as a speedup column.

Step 7. We define the new partition \( \bar{P} \) and the adjustment function \( \bar{\alpha} \) on each of the speedup columns just created: Our model name \( F \) has the form \( F = \{(p_s, g_s') \}_{s \in |F|} \subseteq (P \times G)^{|F|} \). In this speedup column we have the particular speedup orbit, taken from the orbit of \( y' \), which is a sequence \( \{(\bar{x}_s, g_s)\}_{s \in |F|} \) of points in \( \bar{X} \times G \). We define \( \bar{P}_1 \) on the points \( \{\bar{x}_s\}_{s \in |F|} \) by setting, for each \( s \), \( \bar{P}_1 (\bar{x}_s) = p_s \), and we make \( \bar{P}_1 \) constant on the column level that contains \( \bar{x}_s \). (We can view \( \bar{P} \) as either a partition of \( \bar{X} \) or of \( \bar{X} \times G \) without confusion). We define \( \bar{\alpha} \) on the points \( \{\bar{x}_s\}_{s \in |F|} \) by requiring, for each \( s \), \( \bar{\alpha} (\bar{x}_s) g_s = g'_s \). We extend \( \bar{\alpha} \) to the whole speedup column we have constructed as follows. Let \( L \) denote the base of this column. For each point \( \bar{\varepsilon} \in \bar{X} \) such that \( \{\bar{\varepsilon}\} \times G \subset L \), we consider the orbit of \( (\bar{\varepsilon}, g_0) \) under the speedup \( \bar{S}_1 \) that we have defined (so far just on this column). The \( G' \)-coordinates of the points on this orbit are rotations of \( g_0 \) by successive products of the skewing function \( \sigma \) associated with \( \bar{S}_1 \). Namely, \( c (\bar{S}_1 (\bar{\varepsilon}, g_0)) = (\bar{T}_1 (\bar{\varepsilon}), \sigma (\bar{\varepsilon}) (\bar{\varepsilon}) g_0) \)

where \( \sigma (\bar{\varepsilon}) = \sigma (\bar{T}_1 (\bar{\varepsilon})) \ldots \sigma (\bar{T}_1^t (\bar{\varepsilon})) \sigma (\bar{T}_1 (\bar{\varepsilon})) \sigma (\bar{\varepsilon}) \). But for each \( j \in [t] \),

\[ \rho (\bar{\sigma} (\bar{T}_1^j (\bar{\varepsilon})), \bar{\sigma} (\bar{T}_1^j (\bar{x}_0))) < \zeta' \]

and \( t \leq M' \), so by the choice of \( \zeta' \) (see \([3, 12]\)), we get \( \rho (\sigma (\bar{\varepsilon}), \sigma (\bar{\varepsilon}) (\bar{x}_0)) < \zeta \). We define \( \bar{\alpha} \) on the \( \bar{T}_1 \) orbit of \( \bar{\varepsilon} \) by requiring that \( \bar{\alpha} (\bar{T}_1 (\bar{\varepsilon})) \sigma (\bar{\varepsilon}) (\bar{\varepsilon}) g_0 = g'_t \)

In other words, the “adjusted” \( G \)-coordinates of the points on this orbit are identical to the \( G \)-coordinates of the corresponding terms in the model name \( F \).

The useful blocks fill at least a \( (1 - \delta) \)-fraction of each window, and we used a \( (1 - \frac{\delta}{100}) \)-fraction of the useful blocks in each window. We may assume that \( \delta_1 \) was chosen so that \( (1 - \frac{\delta_1}{100}) (1 - \delta) > (1 - 2\delta) \). Since the windows \( W_t \) (even those not within \( p \) windows of the top and bottom) occupy at least a \( (1 - \zeta) \)-fraction of
for every point $z$. We note that the partition $\bar{\lambda}$ also extends the definition of $\bar{\lambda}$. Condition (3.13) is met if the following conditions on $(\bar{\lambda})$:

Next we establish (3.3) and (3.4). Fix a column $\bar{P}_1$ and $\bar{\alpha}_1$ to the complement of $\bar{\lambda}$, by including the complement of $\bar{\lambda}$ in a single atom of $\bar{P}_1$ and by setting $\bar{\alpha}_1$ equal to $\text{id}_G$ on the complement of $\bar{\lambda}$.

We now verify that the conclusions of the lemma hold.

To establish (3.2), the only property requiring explanation is the fourth of the definition of $(n_1, \delta_1)$—regularity. But for all $z$ as in step 9, properties (a.) and (b.) in the formation of $F$ imply that, with $W = S_{i_{\in [n_1]}}(z)$,

$$\|\text{dist}_{W_{1}} \bigvee_{i \in [n_1]} \bar{S}_{i}^{-1} (\bar{P}_1 \lor \bar{\alpha}_1), \text{dist}_{\text{Dom}(\bar{S}_{i}^{n_1})} \bigvee_{i \in [n_1]} \bar{S}_{i}^{-1} (\bar{P} \lor \bar{\alpha}_1) \|_M \leq \frac{\delta_1}{50},$$

And so for each $h \in G$,

$$\|\text{dist}_{W_{1h}} \bigvee_{i \in [n_1]} \bar{S}_{i}^{-1} (\bar{P}_1 \lor \bar{\alpha}_1), \text{dist}_{\text{Dom}(\bar{S}_{i}^{n_1})} \bigvee_{i \in [n_1]} \bar{S}_{i}^{-1} (\bar{P} \lor \bar{\alpha}_1) \|_M \leq \frac{\delta_1}{50}.$$
(b) \(\rho' \left[ F_n (\psi (i)) , N \cup u (P \vee \tilde{c}) (\tau_{0, t} \psi (i)) \right] < \delta + \frac{\delta}{100} \)
and

(c) \(\rho' \left[ N \cup u (P \vee \tilde{c}) (\tau_{0, t} \psi (i)) ; N \cup u (P \vee \tilde{c}) (\tau_{P, P + t + i, t} (i)) \right] < \zeta \).

(Not that \(\tau_{P, P + t + i, t} (i) = \tilde{g}_{j}^{l, t} (i)\).

Each of these conditions holds for a fraction of the set of real 4-tuples which is greater than

(a) \((1 - \zeta) (b) \left( 1 - \frac{\delta}{1 - \delta - \alpha} \right) (c) (1 - 2\zeta)\).

Concerning (b), we know that for all \(y\) in a set \(\nu' \subset \nu \) with \(\frac{|\nu'|}{|\nu|} > (1 - \delta)\) we have \(\rho' \left[ N \cup u (P \vee \tilde{c}) (\tau_{0, t} \psi (i))\right] < \delta\), so if \(\tau_{0, t} (\psi (i)) = y \in \nu'\) then \(\rho' \left[ F_n (\psi (i)) ; N \cup u (P \vee \tilde{c}) (\tau_{0, t} \psi (i))\right] < \rho' \left[ F_n (\psi (i)) , N \cup u (P \vee \tilde{c}) (\tau_{0, t} \psi (i))\right] < \frac{\delta}{100} + \delta\). Here the fact that \(Q \left( F_n (\psi (i)) \right) = Q \left( F_n (\theta y) \right)\) implies \(\rho' \left[ F_n (\psi (i)) , N \cup u (P \vee \tilde{c}) (\tau_{0, t} \psi (i))\right] < \frac{\delta}{100}\).

(b) \((1 - \zeta) (b) \left( 1 - \frac{\delta}{1 - \delta - \alpha} \right) (c) (1 - 2\zeta)\).

Thus, (if \(\zeta\) and \(\zeta\) are sufficiently small with respect to \(\delta\)) for a set of 4-tuples of density greater than \((1 - 2\delta)\), we have condition \((3.13)\).

But the blocks that arise from these 4-tuples (that is the useful blocks with initial points \(\tilde{g}_{j}^{l, t} (i)\)) occupy at least a \((1 - \frac{1}{100})\) fraction of the set of all useful blocks in \(S^{[M]} \psi_{i}\)' and the useful blocks in \(C'\) occupy at least a \((1 - \delta - \zeta)\) fraction of the points in the windows of \(S^{[M]} \psi_{i}\)' and the windows occupy at least a \((1 - \zeta)\) fraction of \(S^{[M]} \psi_{i}\)'. So (if \(\delta_1\) and \(\zeta\) are sufficiently small) the set of column blocks that arise from these 4-tuples occupy at least a \((1 - 2\delta)\) fraction of \(C'\). Since \(C'\) is an arbitrary column and \(\tilde{\mu} \times \lambda (R') > 1 - \zeta\), we obtain \((3.3)\) and \((3.4)\).

To establish \((3.5)\) we note that \(\tilde{\mu} \times \lambda (R \setminus R') < \zeta\) and if \(E\) denotes the set of points in \(R \cap R'\) whose ladder block is not used in the construction of \(\tilde{R}_0\), then \(\tilde{\mu} \times \lambda (E) < \frac{\delta}{100}\) since all the ladder blocks that were used in the construction of \(\tilde{R}_0\) were unbroken, condition \((3.5)\) is obtained.

To establish \((3.6)\) we argue as in the case of condition \((3.2)\). For all points \(z \in B_i\) whose \(G\) - coordinate is the same as that of the initial term in \(F\),

\[
\left\| dist \tilde{g}_{[M] - n + 1,i} \cup i \cup n \right\| \left. \left. \sum_{i \in [n]} \tilde{S}^{-i}_{i} \left( P_{1} \vee \tilde{c} \right) \right) \cup i \cup n \sum_{i \in [n]} \tilde{S}^{-i} (P \vee c) \right\|_{M} < \frac{\delta}{100}.
\]

and for each \(h \in G\),

\[
\left\| dist \tilde{g}_{[M] - n + 1,i} \cup i \cup n \right\| \left. \left. \sum_{i \in [n]} \tilde{S}^{-i}_{i} \left( P_{1} \vee \tilde{c} \right) \right) \cup i \cup n \sum_{i \in [n]} \tilde{S}^{-i} (P \vee c) \right\|_{M} < \frac{\delta}{100}.
\]

and this is sufficient to imply \((3.6)\).

Finally we verify that condition \((3.7)\) holds on a sufficiently large set. Fix a column \(C'\) of \(R'\). Let \(D\) denote the set of real 4-tuples \((l, j, i, t)\) associated with \(C'\) such that:

(a) \(\rho' \left[ F_n (i) , F_n (\psi (i)) \right] < \zeta \),

(b) \(\rho' \left[ F_n (\psi (i)) ; N \cup u (P \vee \tilde{c}) (\tau_{0, t} \psi (i)) \right] < \delta + \frac{\delta}{100} \),

(c) \(\left. \left. \sum_{i \in [n]} S^{-i} (P \vee \tilde{c} \cup 1) \right) \cup i \cup n \sum_{i \in [n]} S^{-i} (P \vee \tilde{c} \cup 1) \right) \left( \tilde{g}_{j}^{l, t} (i) \right) \are \zeta - close.\)
(d) $\tau_{0,\psi}(i)$ is not in one of the miscellaneous atoms of $(Q_0 \lor \mathcal{R}_0)$ constructed in step 1.

As we argued above, conditions (a) and (b) hold for a fraction of real 4-tuples greater than $(1 - \zeta)$ and \(\left(1 - \frac{\delta}{1 - \frac{\zeta}{100}}\right) > 1 - 2\delta\), respectively. By similar arguments, condition (c) holds for a fraction greater than $(1 - 2\zeta)$ and (d) holds for a fraction greater than \(\left(1 - \frac{\epsilon'/100}{1 - \frac{\zeta}{100}}\right) > \left(1 - \frac{\epsilon'}{\zeta}\right)\). Thus all four conditions hold for a fraction greater than \(\left(1 - \frac{\epsilon'}{\zeta}\right)\). (\(\delta, \delta_1\) and \(\zeta\) must have been chosen to make this last inequality hold).

In the present argument we can ignore the partitions $P$ and $\bar{P}$ in the above conditions. In particular, when $(t,l,j,i) \in D$ and writing $c_n(i)$ for the second component of $F_n(i)$, we have that

\[(e)\] $c_n(i)$ and $\bigvee_{u \in [n]} \bar{S}^{-u} \bar{c}(\hat{y}^{l,t}_j(i))$ are uniformly close to within $2\delta$.

We also note that condition (c) gives

\[(f)\] $\bigvee_{u \in [n]} \bar{S}^{-u} (1_{A_1}) (\tau_{0,\psi}(i)) = \bigvee_{u \in [n]} \bar{S}^{-u} (1_{A_1}) \left(\hat{y}^{l,t}_j(i)\right)$.

For each initial triple $(t,l,j)$ (i.e. the initial triple of a real 4-tuple associated with $C'$), and for each $k \in \left[\frac{|\psi|}{n}\right]$, let $\hat{y}^{l,t}_{j,k}$ denote the restriction of $\hat{y}^{l,t}_j$ to the interval $[n_1]+kn_1$. For a given triple $(t,l,j)$ and $k$ we consider $\mathcal{J}_k = \{i \in \mathcal{J} \mid (t,l,j,i) \in D \}$. (We refrain from writing $\hat{J}^{l,t}_{j,k}$ for $\mathcal{J}_k$).

Let $E = \left\{ (t,l,j,k) \mid \left| \mathcal{J}_k \right| > 1 - \sqrt{\frac{2}{|n|}} \right\}$. Then $E$ has density greater than \(\left(1 - \sqrt{\frac{\epsilon'}{\zeta}}\right)\) in the set of all $(t,l,j,k)$.

Fix $(t,l,j,k) \in E$. The range of $\hat{y}^{l,t}_{j,k}$ is an orbit of $\bar{S}_1$ whose initial point we denote by $y$. Let $C'(y)$ denote the level of $C'$ containing $y$. Then $C'(y)$ is the base of an $S_1 - n_1$-column and is a subset of $\Lambda_{n_1}(S_1)$. We will show that (3.7) holds for all points in $C'(y)$.

First we consider the point $y$ itself. If it were the case that $\mathcal{J}_k = \hat{J}_k$, then for all $i \in \mathcal{J}_k$ condition (e) would hold. In addition, if for all $q \in Q$ we let $\mathcal{J}_{k,q} = \{i \in \mathcal{J}_k \mid F_n\psi(i) \in q\}$, we would have

\[
\left\| \text{dist}_{\mathcal{J}_{k,q}} \bigvee_{u \in [n]} \bar{S}^{-u} (1_{A_1}) \left(\hat{y}^{l,t}_j(i)\right), \text{dist}_{\mathcal{J} \cap \psi_0} \bigvee_{u \in [n]} \bar{S}^{-u} (1_{A_1}) \right\|_{\mathcal{M}} < \frac{\epsilon'}{50}
\]

(since the miscellaneous atoms occupy less than an $\frac{\epsilon'}{100}$-fraction of $\psi_0$). Therefore, we would have

\[
\left\| \text{dist}_{\mathcal{J}_k} \bigvee_{u \in [n]} \bar{S}^{-u} (\bar{c} \lor 1_{A_1}) \left(\hat{y}^{l,t}_j(i)\right), \text{dist}_{\mathcal{J} \cap \psi_0} \bigvee_{u \in [n]} \bar{S}^{-u} (\bar{c} \lor 1_{A_1}) \right\|_{\mathcal{M}} < \frac{\epsilon'}{50} + \delta.
\]

(since the $c - n$-names can differ by $\delta$). This would give

\[
\left\| \text{dist}_{\mathcal{J}_k} \bigvee_{u \in [n]} \bar{S}^{-u} (\bar{c} \lor 1_{A_1}) \left(\hat{y}^{l,t}_j(i)\right), \text{dist}_{\mathcal{J} \times G} \bigvee_{u \in [n]} \bar{S}^{-u} (\bar{c} \lor 1_{A_1}) \right\|_{\mathcal{M}} < \frac{\epsilon'}{50} + \delta + \zeta.
\]
But since we only have that $|J_k| > (1 - \sqrt{\frac{\varepsilon'}{20}}) |J_k|$, we have instead
\[
\left\| \text{dist}_J \left( \sum_{u \in [n]} S^{-u} (\bar{e} \vee 1_{A_1}) \left( \bar{g}^{l \cdot t}_j \right), \text{dist}_M \left( \sum_{u \in [n]} S^{-u} (\bar{e} \vee 1_{A_1}) \right) \right\|_M < \frac{\varepsilon'}{50} + \delta + \zeta + \sqrt{\frac{\varepsilon'}{40}} < \sqrt{\frac{\varepsilon'}{30}}
\]
(if $\varepsilon'$, $\delta$, and $\zeta$ are sufficiently small).

For each $i \in J_k$, we write $n(i)$ for the $n$–block beginning at $i$ and $n(J_k)$ for $\bigcup_{i \in J_k} n(i)$. Then we have
\[
\left\| \text{dist}_{n(J_k)} (\bar{e} \vee 1_{A_1}) \left( \bar{g}^{l \cdot t}_j \right), \text{dist}_M (\bar{e} \vee 1_{A_1}) \right\|_M < \sqrt{\frac{\varepsilon'}{30}}.
\]
Since $|n(J_k)| > (1 - \zeta) n_1$ we get
\[
\left\| \text{dist}_{n_1} \left( \sum_{u \in [n_1]} \bar{e} \vee 1_{A_1} \right) \left( \bar{g}^{l \cdot t}_j \right), \text{dist}_M (\bar{e} \vee 1_{A_1}) \right\|_M < \sqrt{\frac{\varepsilon'}{20}}.
\]
In other words,
\[
\left\| \text{dist}_{\bar{S}_1^{[n_1]} y} (\bar{e} \vee 1_{A_1}), \text{dist}_M (\bar{e} \vee 1_{A_1}) \right\|_M < \sqrt{\frac{\varepsilon'}{20}}.
\]
Since $A_1 \in \bar{A}$, $\bar{e}$ is independent of $1_{A_1}$, so for all $h \in G$, $\text{dist}_M (\bar{e} \vee 1_{A_1}) = \text{dist}_M (\bar{e} \vee 1_{A_1})$ and $\text{dist}_{\bar{S}_1^{[n_1]} y} (\bar{e} \vee 1_{A_1}) = \text{dist}_{\bar{S}_1^{[n_1]} y} (\bar{e} \vee 1_{A_1})$, so
\[
\left\| \text{dist}_{\bar{S}_1^{[n_1]} y} (\bar{e} \vee 1_{A_1}), \text{dist}_M (\bar{e} \vee 1_{A_1}) \right\|_M < \sqrt{\frac{\varepsilon'}{20}}.
\]
If $\bar{y} \in C' (y)$ and $\bar{e} (\bar{y}) = \bar{e} (y)$, then for all $u \in [n_1]$, $\rho (\bar{e} (\bar{S}_1^{u} \bar{y}), \bar{e} (\bar{S}_1^{u} y)) < \zeta$ so we get
\[
\left\| \text{dist}_{\bar{S}_1^{[n_1]} y} (\bar{e} \vee 1_{A_1}), \text{dist}_M (\bar{e} \vee 1_{A_1}) \right\|_M < \sqrt{\frac{\varepsilon'}{20}} + \zeta
\]
and so as before, for all $h \in G$
\[
\left\| \text{dist}_{\bar{S}_1^{[n_1]} y} (\bar{e} \vee 1_{A_1}) - \text{dist}_M (\bar{e} \vee 1_{A_1}) \right\|_M < \sqrt{\frac{\varepsilon'}{20}} + \zeta.
\]
For simplicity, let’s suppose that here we got $\varepsilon'$ as an upper estimate, as we could have done. Then for all $z \in C' (y)$ we have
\[
\frac{1}{n_1} \sum_{u \in [n_1]} \bar{1}_{A_1} (\bar{S}_1^{u} z) > \mu (A_1) - \varepsilon'
\]
and conditioning on $\{ u | \bar{S}_1^{u} z \in A_1 \}$, we have
\[
\left\| \text{dist}_{\{ u | \bar{S}_1^{u} z \in A_1 \}} (\bar{e} (\bar{S}_1^{u} z)) - \lambda \right\|_M < \varepsilon'
\]
so that (by the choice of $\varepsilon'$)
\[
\left\| \text{dist}_{\{ u | \bar{S}_1^{u} z \in A_1 \}} (\bar{e} (\bar{S}_1^{u} z)) - \lambda (A_2) \right\|_M < \frac{\varepsilon}{100}
\]
and so
\[
\frac{1}{n_1} \sum_{u \in A} (\bar{S}_1^{u} z) > \mu \times \lambda (A_1) - \varepsilon.
\]
Thus for all $(t, l, j, k)$ as above, and all $z$ in the base of the corresponding $\bar{S}_1 - n_1$–column block, condition (3.7) holds. But the union of these column blocks
covers at least a \((1 - \sqrt{\frac{\delta}{30}})\) -fraction of the portion of \(R_0 \cap C\) and hence at least a \((1 - 4\delta)(1 - \sqrt{\frac{\delta}{30}})\) -fraction of \(R_1 \cap C\). Therefore, on at least a \((1 - \varepsilon)\) -fraction of \(\Lambda_{n_1} (\bar{S}_1) \cap C'\) we get condition 4.17. Since the same argument applies to every other column of \(R\), the argument is complete. □

The above lemma will now be extended to the case of countable partitions. The statement is identical, with the understanding that the partitions \(P\) and \(\bar{P}\) are countably infinite.

**Lemma 16. (Distribution Improvement Lemma for countable partitions)**

**Proof.** Fix \(\varepsilon > 0\) and \(A_2 \subset G\) open. Let \(n\) and \(\delta\) be given by lemma \(^{15}\) with respect to \(\varepsilon\) and \(A_2\). Suppose that \((S, T, \sigma, X)\) and \((\bar{S}_0, \bar{T}_0, \bar{\sigma}_0, \bar{X})\) are ergodic \(G\)-extensions on \((X \times G)\) and \((\bar{X} \times G)\) respectively with partitions \(P\) and \(\bar{P}\) as in the statement of this lemma. Fix \(n_1, \delta_1\) and \(A_1 \subset \bar{X}\). We assume that the elements of \(P\) are indexed by \(\mathbb{N}\), and we let \(P_N\) denote the partition formed by replacing the set of elements of \(P\) indexed by integers greater than \(N\) by their union. Choose \(N\) so large that

\[
\left\| \text{dist}_{X \times G} \bigvee_{i \in [n_1]} S^{-i} (P_N \vee c), \text{dist}_{X \times G} \bigvee_{i \in [n_1]} S^{-i} (P \vee c) \right\|_\mathcal{M} < \frac{\delta_1}{2}
\]

and

\[
\left\| \text{dist}_{X \times G} \bigvee_{i \in [n]} S^{-i} (P_N \vee c), \text{dist}_{\text{Dom}(S^n)} \bigvee_{i \in [n]} \bar{S}^{-i} (\bar{P} \vee c) \right\|_\mathcal{M} < \delta.
\]

Apply lemma \(^{15}\) to the systems \((S, P_N), (\bar{S}, \bar{P}), (\bar{S}_0)\), and the set \(A_1 \times A_2\) but using the parameter \(\frac{\delta_1}{2}\) instead of \(\delta_1\). This gives us a new partial speedup \(\bar{S}_1\) of \(\bar{S}_0\), a partition \(\bar{P}_1\) and a function \(\bar{\alpha}\) satisfying the conclusions of the lemma \(^{15}\) (with respect to \((S, P_N)\)). But then condition 3.14 gives the conclusion 4.16. All the other conclusions don’t refer to \(P\), so they are met as well. □

We note that the order of quantifiers in the statement of lemma \(^{15}\), namely that \(\delta\) and \(n\) depend only on \(\varepsilon\) and \(A_2\), was used in an essential way in this argument.

### 4. Factor Theorem

Our goal here is to obtain the following:

**Theorem 1.** For all \(\varepsilon > 0\) there exists \(\delta > 0\) and \(n \in \mathbb{N}\) such that if \((S, T, \sigma, X)\) and \((\bar{S}_0, \bar{T}_0, \bar{\sigma}_0, \bar{X})\) are ergodic \(G\)-extensions on \(X \times G\) and \(\bar{X} \times G\), respectively and if \(P\) is a generator for \((T, X)\) and \(\bar{P}\) is a partition of \(\bar{X}\), such that

\[
\left\| \text{dist}_{X \times G} \bigvee_{i \in [n]} S^{-i} (P \vee c), \text{dist}_{\bar{X} \times G} \bigvee_{i \in [n]} \bar{S}^{-i} (\bar{P} \vee c) \right\|_\mathcal{M} < \delta.
\]
then there exists an ergodic $G$–speedup $(\hat{S}, \hat{T}, \hat{\sigma}, \hat{X})$ of $(\bar{S}_0, \bar{T}_0, \bar{\sigma}_0, \bar{X})$ and a partition $\hat{P}$ of $\hat{X}$ and a measurable function $\hat{\alpha} : \hat{X} \to G$ such that $|\hat{P} - \hat{P}| < \varepsilon$,

$$\int_{\hat{X}} \rho(\hat{\alpha}(\hat{x}), id_G) \, d\hat{\mu} < \varepsilon,$$

$$\hat{\mu} \times \lambda \{ (\hat{x}, g) \mid \hat{S}(\hat{x}, g) \neq \hat{S}_0(\hat{x}, g) \} < \varepsilon,$$

and for all $n \in \mathbb{N}$,

$$\left\| \text{dist}_{X \times G} \bigvee_{i \in [n]} S^{-i} (P \vee \alpha), \text{dist}_{X \times G} \bigvee_{i \in [n]} \hat{S}^{-i} \left( \hat{P} \vee \hat{\alpha} \right) \right\|_M = 0.$$

In particular, $(\hat{S}, \hat{T}, \hat{\sigma}, \hat{X})$ has $(S, T, \sigma, X)$ as a $G$–factor.

This theorem will follow from the next lemma, which will be proved by repeated application of lemma [16].

**Lemma 17.** For all $\varepsilon > 0$ there exists $\delta > 0$ and $n \in \mathbb{N}$ so that if $(S, T, \sigma, X)$ and $(\bar{S}_0, \bar{T}_0, \bar{\sigma}_0, \bar{X})$ are ergodic $G$–extensions on $X \times G$ and $\bar{X} \times G$, respectively, and $(\hat{S}, \hat{T}, \hat{\sigma}, \hat{X})$ is a partial $G$–speedup of $\bar{S}_0$, and $P$ and $\bar{P}$ are partitions of $X$ and $\bar{X}$, respectively, such that $(\hat{S}, \hat{P})$ is $(n, \delta)$–regular, and

$$\left\| \text{dist}_{X \times G} \bigvee_{i \in [n]} S^{-i} (P \vee \alpha), \text{dist}_{\text{Dom}(\bar{S}_0)} \bigvee_{i \in [n]} \bar{S}^{-i} \left( \bar{P} \vee \bar{\alpha} \right) \right\|_M < \delta,$$

then there is an ergodic $G$–speedup $(\hat{S}, \hat{T}, \hat{\sigma}, \hat{X})$ of $(\bar{S}_0, \bar{T}_0, \bar{\sigma}_0, \bar{X})$ and a partition $\hat{P}$ of $\hat{X}$ and a measurable function $\hat{\alpha} : \hat{X} \to G$ such that

$$|\hat{P} - \hat{P}| < \varepsilon,$$

$$\int_{\hat{X}} \rho(\hat{\alpha}(\hat{x}), id_G) \, d\hat{\mu} < \varepsilon,$$

$$\hat{\mu} \times \lambda \{ (\hat{x}, g) \mid \hat{S}(\hat{x}, g) \neq \hat{S}_0(\hat{x}, g) \} < \varepsilon,$$

and for all $n \in \mathbb{N}$,

$$\left\| \text{dist}_{X \times G} \bigvee_{i \in [n]} S^{-i} (P \vee \alpha), \text{dist}_{\text{Dom}(\bar{S}_0)} \bigvee_{i \in [n]} \bar{S}^{-i} \left( \bar{P} \vee \bar{\alpha} \right) \right\|_M = 0.$$

**Proof.** Fix $\varepsilon > 0$. Choose $\varepsilon_k > 0$ so that $\sum_{k=0}^{\infty} \varepsilon_k < \frac{\varepsilon}{2}$. Fix a sequence of measurable rectangles $\left\{ A^{(k)} \right\} = \bigcup_{k=0}^{\infty} A^{(k)} \times A^{(k)} \in \mathcal{X} \times G$, where each $A^{(k)}$ is open, each rectangle appears infinitely often in the sequence, and the sequence is dense in the measure algebra of $(\hat{X} \times G, \hat{\mu} \times \lambda)$. For each $k$, choose $\delta_k < \frac{\varepsilon_k}{2}$ and $n_k$ by applying lemma [16] with respect to $\varepsilon_k$ and $A^{(k)}$. We will see that $\delta_0$ and $n_0$ serve as the $\delta$ and $n$ in the conclusion of this theorem.

Suppose now that $(S, T, \sigma, X)$ and $(\bar{S}_0, \bar{T}_0, \bar{\sigma}_0, \bar{X})$ are given as in the statement of the lemma. Let $\hat{S}_1, \hat{P}_1$ and $\hat{\sigma}_1$ be the partial speedup of $\bar{S}_0$, the partition and
the function given by lemma 16. Here we use $A^{(0)} = A_1^{(0)} \times A_2^{(0)}$ as the rectangle in the statement of lemma 16. Roughly speaking, lemma 16 allows us to make an $\varepsilon$-small modification of $(\bar{S}, \bar{P})$ to obtain $(\bar{S}_1, \bar{P}_1)$. The conclusions of lemma 16 and the choice of $\delta_1$ and $n_1$ allow us to apply lemma 16 again to the new $G$ extension $\bar{S}_{1}^{\alpha_1}$ and its partial speedup $\bar{S}_1^{\alpha_1}$ to obtain $\bar{S}_2, \bar{P}_2$ and $\alpha_2$ which meet the conclusions of lemma 16 with respect to the rectangle $A^{(1)} = A_1^{(1)} \times A_2^{(1)}$. That is, we make an $\varepsilon_1$-small modification of $(\bar{S}_1^{\alpha_1}, \bar{P}_1)$ and obtain a new partial speedup $\bar{S}_2$ of $\bar{S}_0$, a partition $\bar{P}_2$ and a function $\alpha_2$. At this point, we let $\beta_2 = \alpha_2 \alpha_1$ so that the $G$ extension $\bar{S}_2$ and its partial speedup $\bar{S}_2^\beta$ meet the conditions to which lemma 16 can be applied once again.

Continuing in this way we obtain a sequence of speedups $\bar{S}_k$, partitions $\bar{P}_k$ and functions $\alpha_k$ so that, writing $\bar{\alpha}_k = \prod_{j=0}^{k-1} \alpha_{n_j}$, we have for each $k$,

\begin{equation}
|\bar{P}_{k+1} - \bar{P}_k| < \varepsilon_k,
\end{equation}

\begin{equation}
\int_X \rho(\bar{\alpha}_k(x), id_G) d\bar{\mu} < \varepsilon_k,
\end{equation}

if $D_k$ denotes the set of points in the speedup tower of $\bar{S}_k$ such whose ladder block is broken by $\bar{S}_{k+1}$, then

\begin{equation}
\bar{\mu} \times \lambda(D_k) < \varepsilon_k,
\end{equation}

\begin{equation}
\left\| \text{dist}_{X \times G} \bigvee_{i \in [n_k]} S^{-i} (P \cup c), \text{dist}_{\text{Dom}(\bar{S}_k)} \bigvee_{i \in [n_k]} \bar{S}_k^{-i} (\bar{P}_k \cup \bar{\beta}_k c) \right\|_\mathcal{M} < \delta_k,
\end{equation}

and the set of $y \in \Lambda_{n_k}(\bar{S}_k)$ such that

\begin{equation}
\frac{1}{n_k} \sum_{i \in [n_k]} \chi_{A^{(k)}} (\bar{S}_k(y)) > (\bar{\mu} \times \lambda) \left( A^{(k)} \right) - \varepsilon_k
\end{equation}

has measure greater than $(1 - \varepsilon_k) (\bar{\mu} \times \lambda) \left( \Lambda_{n_k} (\bar{S}_k) \right)$. 

(Recall that $\Lambda_{n_k}(\bar{S}_k)$ denotes the $n_k$-ladder in the speedup tower of $\bar{S}_k$).

Conditions 4.0 and 4.3 and the fact that, for each $k$, $\bar{\mu} \times \lambda (\text{Dom}(\bar{S}_k)) > 1 - \varepsilon_k$ imply that there is a partition $\bar{P}$ and a function $\bar{\alpha}$ such that $\lim_{k \to \infty} |\bar{P}_k - \bar{P}| = 0$ and $\lim_{k \to \infty} \bar{\beta}_k = \bar{\alpha}$ a.e., where $\bar{P}$ and $\bar{\alpha}$ satisfy conditions 4.2 and 4.3.

Condition 4.3 implies (again using $\bar{\mu} \times \lambda (\text{Dom}(\bar{S}_k)) > 1 - \varepsilon_k$) that the partial transformations $\bar{S}_k$ converge almost everywhere to a transformation $\bar{S}$ that is a $G$-speedup of $\bar{S}_0$, and $\bar{S}$ satisfies 4.3.

To establish condition 4.5 we fix $n'$, and $\delta'$ and choose $k$ so that $n_k > n'$. We know that

\begin{equation}
\left\| \text{dist}_{X \times G} \bigvee_{i \in [n_k]} S^{-i} (P \cup c), \text{dist}_{\text{Dom}(\bar{S}_k)} \bigvee_{i \in [n_k]} \bar{S}_k^{-i} (\bar{P}_k \cup \bar{\beta}_k c) \right\|_\mathcal{M} < \delta_k
\end{equation}
and so the same is true for the \(n'\) distribution:

\[
\left\| \text{dist}_{X \times G} \bigvee_{i \in [n']} S^{-i} (P \lor c), \text{dist}_{\text{Dom}(S^n)} \bigvee_{i \in [n']} \tilde{S}^{-i}_k (\tilde{P} \lor \beta_k c) \right\|_M < \delta_k
\]

Moreover, the set of \(\tilde{S}_k\)–ladder blocks that are broken by \(\tilde{S}\) has measure less than \(\sum_{i=k}^{\infty} \varepsilon_i\), and \(\tilde{P} - \tilde{P} < \sum_{i=k}^{\infty} \varepsilon_i\), and \(\int_X \rho (\tilde{\beta} (\tilde{x}), \alpha (\tilde{x})) d\mu < \sum_{i=k}^{\infty} \varepsilon_i\). Since the measure of the speedup tower for \(\tilde{S}_k\) is greater than \(1 - \frac{\varepsilon_k}{2}\), we see that if \(k\) is sufficiently large, (so that the set of points whose \(\tilde{S}_k - n\)–orbits are not wholly contained in a ladder block for \(\tilde{S}_k\) is small), we get

\[
\left\| \text{dist}_{X \times G} \bigvee_{i \in [n']} S^{-i} (P \lor c), \text{dist}_{\tilde{X} \times \tilde{G}} \bigvee_{i \in [n']} \tilde{S}^{-i} (\tilde{P} \lor \alpha c) \right\|_M < \delta'
\]

Since this is true for all \(n'\) and \(\delta'\), we have condition 4.5.

Finally, we show that \(\hat{S}\) is ergodic. Fix rectangles \(A^{(i)}\) and \(A^{(j)}\) where

\[
(\hat{\mu} \times \lambda) \left( A^{(i)} \right) < (\hat{\mu} \times \lambda) \left( A^{(j)} \right).
\]

Condition 4.10 and the fact that the measure of the speedup tower for \(\tilde{S}_k\) is greater than \(1 - \frac{\varepsilon_k}{2}\) implies that for all \(\varepsilon'\) there exists \(k\) such that the set of \(\tilde{S}_k\)–ladder blocks on which \(A^{(i)}\) has density within \(\varepsilon'\) of the measure of \(A^{(j)}\) exceeds \(1 - \varepsilon'\). For \(l > k\), most \(\tilde{S}_l\)–ladder blocks are mostly covered by these \(\tilde{S}_k\)–ladder blocks, so we get the stronger fact that for all \(\varepsilon'\) and for all sufficiently large \(k\), the set of \(\tilde{S}_k\)–ladder blocks on which \(A^{(i)}\) has density within \(\varepsilon'\) of the measure of \(A^{(j)}\) exceeds \(1 - \varepsilon'\). Applying this to both \(A^{(i)}\) and \(A^{(j)}\), we can choose \(k\) so that the above condition holds for both rectangles, and in addition, the set of \(\tilde{S}_k\)–ladder blocks that are broken by \(\hat{S}\) has measure less than \(\varepsilon'\). If \(\varepsilon'\) is small enough, we conclude that there is a transformation \(S'\) in the full group of \(\hat{S}\) so that \((\hat{\mu} \times \lambda) \left( S' (A^{(i)} \cap A^{(j)}) \right) > (1 - \varepsilon') (\hat{\mu} \times \lambda) \left( A^{(i)} \right)\).

We conclude from lemma 4.14 that \(\hat{S}\) is ergodic. \(\square\)

We now give the proof of theorem 4.1 using lemma 4.17.

**Proof.** (of theorem 4.1) Fix \(\varepsilon > 0\). Choose \(\delta\) and \(n\) by lemma 4.12 with respect to \(\hat{S}\). Suppose that \((S,T,\sigma,X)\) and \((\tilde{S}_0, \tilde{T}_0, \tilde{\sigma}, \tilde{X})\) are ergodic \(G\)–extensions and \(P\) and \(\tilde{P}\) are partitions satisfying the hypotheses of theorem 4.1 but where the distribution match is to within \(\frac{\delta}{2}\). Lemma 4.1 gives the following: For all \(\zeta > 0\) there exists \(L (\zeta) \in \mathbb{N}\) so that

1. for all \(L \geq L (\zeta)\), \((1 - \zeta)\)–most points \(\tilde{x} \in \tilde{X}\) have the property that for all \(g \in G\),

\[
\left\| \text{dist}_{\tilde{S}^{L L (\zeta)}} \bigvee_{i \in [n]} \tilde{S}^{-i}_0 (\tilde{P} \lor c), \text{dist}_{\tilde{X} \times \tilde{G}} \bigvee_{i \in [n]} \tilde{S}^{-i}_0 (\tilde{P} \lor c) \right\|_M < \zeta
\]

and

2. the interval \([L - 1]\) can be \((1 - \zeta)\)–disjointly covered by a set of intervals of
and if (5.1) use it to obtain theorem 2.

Consequently, lemma 17 gives us an ergodic speedup where these groups can be chosen to be as long as we please. Consequently, for each such $\bar{x}$ we can speed up the $T - L$-orbit of $\bar{x}$ (and correspondingly speed up the $S_0 - L$-orbit of $(\bar{x}, g)$, for each $g \in G$) by skipping over any points that are not in the orbit blocks chosen by these intervals. If the lengths of the consecutive groups of $n$ blocks are sufficiently large compared to $n$, then the distribution of $(\bar{S}_0, P \lor c) - n$-names on such a $G$-speedup orbit segment will be $2\zeta$ close to the distribution of $(\bar{S}_0, P \lor c) - n$-names on $\bar{X} \times G$.

Now choose a Rokhlin tower for $\bar{S}_0$, measurable with respect to $\bar{X}$ and of height $L$ so that all points in its base are of the above type. For each point $\bar{x}$ of the base, implement the speedup described above, and remove just enough levels from the top of the orbit segment above $\bar{x}$ so that the tower which remains is of constant height $L'$, where $L'$ is a multiple of $n$. If $\zeta$ was chosen sufficiently small, this gives a $G$-speedup $\bar{S}$ and a $(\delta, n)$-regular speedup tower that satisfy the hypotheses of lemma 17. In addition, we may arrange that

$$\bar{\mu} \times \lambda \{ (\bar{x}, g) \mid \bar{S}(\bar{x}, g) \neq \bar{S}_0(\bar{x}, g) \} < \frac{\varepsilon}{2}$$

Consequently, lemma 17 gives us an ergodic speedup $\left(\bar{S}, \bar{T}, \bar{\sigma}, \bar{X}\right)$ satisfying the conclusions of theorem 1. □

5. Isomorphism theorem

We now wish to prove our main theorem:

**Theorem 2.** Let $(S, T, \sigma, X)$ and $(\bar{S}_0, \bar{T}_0, \bar{\sigma}_0, \bar{X})$ be ergodic $G$-extensions on $X \times G$ and $\bar{X} \times G$, respectively. Then for all $\varepsilon > 0$ there exists an ergodic $G$-speedup $(\bar{S}, \bar{T}, \bar{\sigma}, \bar{X})$ of $(\bar{S}_0, \bar{T}_0, \bar{\sigma}_0, \bar{X})$ such that

$$\bar{\mu} \times \lambda \{ (\bar{x}, g) \mid \bar{S}(\bar{x}, g) \neq \bar{S}_0(\bar{x}, g) \} < \varepsilon,$$

and $(\bar{S}, \bar{T}, \bar{\sigma}, \bar{X})$ and $(S, T, \sigma, X)$ are $G$-isomorphic.

We will first prove a version of this theorem analogous to theorem 1 and then use it to obtain theorem 2.

**Theorem 3.** For all $\varepsilon > 0$ there exists $\delta > 0$ and $n \in \mathbb{N}$ such that if $(S, T, \sigma, X)$ and $(\bar{S}_0, \bar{T}_0, \bar{\sigma}_0, \bar{X})$ are ergodic $G$-extensions on $X \times G$ and $\bar{X} \times G$, respectively and if $P$ is a generator for $(T, X)$ and $\bar{P}$ is a partition of $\bar{X}$, such that

$$\left\| \text{dist}_{X \times G} \bigcup_{i \in [n]} S^{-i}(P \lor c), \text{dist}_{\bar{X} \times G} \bigcup_{i \in [n]} \bar{S}_0^{-i}(\bar{P} \lor c) \right\| < \delta,$$
then there exists an ergodic $G$–speedup $\left(\hat{S}, \hat{T}, \hat{\sigma}, \hat{X}\right)$ of $\left(S_0, T_0, \sigma_0, X\right)$ and a generator $\hat{P}$ for $\left(\hat{T}, \hat{X}\right)$ and a measurable function $\hat{\alpha} : \hat{X} \to G$ such that $|\hat{P} - \hat{P}| < \varepsilon$, 

$$\int_{\hat{X}} \rho(\hat{\alpha}(x), id_G) \, d\hat{\mu} < \varepsilon,$$

and for all $n \in \mathbb{N}$,

$$\left\| \text{dist}_{X \times G} \bigwedge_{i \in [n]} S^{-i}(P \vee c) \bigwedge_{i \in [n]} S^{-i}(\hat{P} \vee \hat{\alpha} \vee c) \right\|_M = 0.$$

**Lemma 18.** Suppose that $\left(S, T, \sigma, X\right)$ and $\left(S_0, T_0, \sigma_0, X\right)$ are ergodic $G$–extensions on $X \times G$ and $\hat{X} \times G$, respectively, where $(S, T, \sigma, X)$ is a $G$–factor of $(\hat{S}_0, T_0, \sigma_0, \hat{X})$ via a factor map $\Phi$ of the form $\Phi(x, g) = (\phi(x), g)$. Suppose that $P$ is a partition of $X$ and $\hat{P} = \varphi^{-1}(P)$. Let $\hat{Q}$ be a partition of $\hat{X}$. Then for all $\zeta > 0$ and $n \in \mathbb{N}$ there is a partition $Q$ of $X$ such that

$$\left(\hat{S}, \hat{T}, \hat{\sigma}, \hat{X}\right) \text{ and } (S, T, \sigma, X) \text{ are } G\text{–isomorphic.}$$

**Proof.** Choose $n_1 > n$ and construct a Rokhlin tower $\tau$ of height $n_1$ for $S$, measurable with respect to $X$. Let $\bar{\tau} = \Phi^{-1}(\tau)$. Choose $\zeta_1 < \zeta$ and divide each $(T, P)$ column in $\tau$ into finitely many subcolumns on which the values of $\sigma$ form a set of diameter less than $\zeta_1$. If $C$ is such a column, then $\Phi^{-1}(C)$ is a column of $\bar{\tau}$ with the same property. We divide $\Phi^{-1}(C)$ further into subcolumns on each of whose levels $Q$ is constant. Then we divide $C$ into a set of subcolumns with the same conditional distribution, and we define $Q$ to give each the $Q - n_1$–name of the subcolumn of $\Phi^{-1}(C)$ that it is associated with.

If $n_1$ is chosen big enough, and $\zeta_1$ small enough, then condition 5.2 is obtained. \[\square\]

**Proof.** (of theorem 3) Fix $\varepsilon > 0$ and choose $\delta$ and $n$ as in theorem 1 with respect to $S_0$. Suppose that $(S, T, \sigma, X)$ and $(\hat{S}_0, T_0, \sigma_0, \hat{X})$ are ergodic $G$–extensions on $X \times G$ and $\hat{X} \times G$, respectively and if $P$ is a finite generator for $(T, X)$ and $\hat{P}$ is a partition of $\hat{X}$, such that

$$\left\| \text{dist}_{X \times G} \bigwedge_{i \in [n]} S^{-i}(P \vee c) \bigwedge_{i \in [n]} \hat{S}^{-i}(\hat{P} \vee c) \right\|_M < \delta,$$

By theorem 1 there is an ergodic $G$–speedup $\hat{S}_1$ of $\hat{S}_0$ and a partition $\hat{P}_1$ of $\hat{X}$ and a measurable function $\hat{\alpha}_1 : \hat{X} \to G$ such that $|\hat{P} - \hat{P}_1| < \frac{\varepsilon}{2}$,

$$\int_{\hat{X}} \rho(\hat{\alpha}_1(x), id_G) \, d\hat{\mu} < \frac{\varepsilon}{2},$$

$$\hat{\mu} \times \lambda \{(x, g) \mid \hat{S}_1(x, g) \neq \hat{S}_0(x, g)\} < \frac{\varepsilon}{2}.$$
and for all \( n \in \mathbb{N} \),
\[
\left\| \frac{\text{dist}_{X \times G} \bigvee_{i \in [n]} S^{-i} (P \vee c), \text{dist}_{ar{X} \times G} \bigvee_{i \in [n]} \left( \bar{S}_{1}^{\bar{a}_{1}} \right)^{-i} (\bar{P}_{1} \vee c) }{M} \right\| = 0.
\]
The last condition says that the \( G \)-extension \( \bar{S}_{1}^{\bar{a}_{1}} \) has \( S \) as a \( G \)-factor, via a factor map which is the identity on the \( G \)-coordinate, and which has \( \bar{P}_{1} \) as the preimage of \( P \).

Fix a sequence \( \{ \varepsilon_{i} \}_{i=1}^{\infty} \) so that \( \sum_{i=1}^{\infty} \varepsilon_{i} < \frac{1}{2} \) and a sequence \( \{ \bar{A}_{i} \}_{i=1}^{\infty} \) of sets in \( \bar{X} \) that are dense in the measure algebra of \( \bar{X} \) and in which each of these sets appears infinitely often. Choose \( \delta_{1} \) and \( n_{1} \) by theorem 1 with respect to \( \varepsilon_{1} \) and let \( A_{1} \subset X \) be chosen (using lemma 15) so that,
\[
\left\| \frac{\text{dist}_{X \times G} \bigvee_{i \in [n_{1}]} S^{-i} (P \vee 1_{A_{1}} \vee c), \text{dist}_{\bar{X} \times G} \bigvee_{i \in [n_{1}]} \left( \bar{S}_{1}^{\bar{a}_{1}} \right)^{-i} (\bar{P}_{1} \vee 1_{A_{1}} \vee c) }{M} \right\| < \delta_{1}.
\]

Applying theorem 1 again we get an ergodic \( G \)-speedup \( \bar{S}_{2} \) of \( \bar{S}_{1} \) (and hence of \( S_{0} \)) and a partition \( P_{2} \vee 1_{\bar{A}_{1}} \) of \( \bar{X} \) and a function \( \bar{\alpha}_{2} : \bar{X} \to G \) such that
\[
\left| \bar{P}_{1} \vee 1_{\bar{A}_{1}} - P_{2} \vee 1_{\bar{A}_{1}} \right| < \varepsilon_{1},
\]
\[
\int_{\bar{X}} \rho (\bar{\alpha}_{2} (\bar{x}), id_{G}) \, d\bar{\mu} < \varepsilon_{1},
\]
\[
\bar{\mu} \times \lambda \{ (\bar{x}, g) \mid \bar{S}_{2} (\bar{x}, g) \neq \bar{S}_{1} (\bar{x}, g) \} < \varepsilon_{1},
\]
and for all \( n \in \mathbb{N} \),
\[
\left\| \frac{\text{dist}_{X \times G} \bigvee_{i \in [n]} S^{-i} (P \vee 1_{A_{1}} \vee c), \text{dist}_{\bar{X} \times G} \bigvee_{i \in [n]} \bar{S}_{2}^{\bar{\alpha}_{2}} (\bar{P}_{2} \vee 1_{A_{1}} \vee c) }{M} \right\| = 0,
\]
where \( \bar{\alpha}_{1} = \bar{\alpha}_{2} \bar{a}_{1} \). In other words, the \( G \)-extension \( \bar{S}_{2}^{\bar{\alpha}_{2}} \) has \( S \) as a \( G \)-factor, via a factor map which is the identity on the \( G \)-coordinate, and which has \( \bar{P}_{2} \vee 1_{A_{1}} \) as the preimage of \( P \vee A_{1} \). Since \( P \) is a generator for \( T \) we have \( A_{1} \subset \bigvee_{i=-\infty}^{\infty} T_{-i}^{-1} (P_{2}) \).

Since \( \left| 1_{A_{1}} - 1_{\bar{A}_{1}} \right| < \varepsilon_{1} \) we know that for some \( m_{1} \) we have
\[
\bar{A}_{1} \subset \bigvee_{i \in [-m_{1}, m_{1}]} T_{-i}^{-1} (P_{2}).
\]

We choose \( \eta_{2} > 0 \) so that for every transformation \( \hat{T} \) of \( \bar{X} \) and partition \( \hat{P} \) of \( \bar{X} \) such that
\[(5.4) \quad \left| \hat{P} - \hat{P}_{1} \right| < \eta_{2}\]
and
\[(5.5) \quad \bar{\mu} \left\{ \bar{x} \in \bar{X} \mid \hat{T} (\bar{x}) \neq \hat{T}_{1} (\bar{x}) \right\} < \eta_{2}\]
we get
\[
\bar{A}_{1} \subset \bigvee_{i \in [-m_{1}, m_{1}]} \hat{T}^{-i} (\hat{P}).
\]
We will continue making successive speedups and partitions, making sure that the limiting process \((\hat{T}, \hat{P})\) satisfies conditions \([\text{A.4}]\) and \([\text{A.14}]\). To proceed, we replace the numbers \(\{\varepsilon_i\}_{i=2}^\infty\) by smaller numbers (also called \(\varepsilon_i\)) so that \(\sum_{i=2}^\infty \varepsilon_i < \eta_2\). We then repeat the above argument, applying it to the partition \(A_2 = \{A_2, \bar{X} \setminus A_2\}\) and \(\varepsilon_2\) and the process \(S_{\varepsilon_2}^2 (\bar{P} \vee c)\).

Continuing in this way we obtain a sequence of speedups \(\bar{S}_k\) and partitions \(\bar{P}_k\) and functions \(\bar{\alpha}_k : \bar{X} \rightarrow G\) and integers \(m_k\) such that, for each \(k\) (and writing \(\bar{\beta}_k = \prod_{j=0}^{k-1} \bar{\alpha}_{k-j}\)), \(|\bar{P}_{k+1} - \bar{P}_k| < \varepsilon_k\),

\[
\int_X \rho (\bar{\alpha}_{k+1} (\bar{x}), id_G) d\bar{\mu} < \varepsilon_k,
\]

\[
\bar{\mu} \times \lambda \{(\bar{x}, g) \in \bar{X} \times G \mid \bar{S}_{k+1}(\bar{x}, g) \neq \bar{S}_k(\bar{x}, g)\} < \varepsilon_k
\]

for all \(n\)

\[
\left\| \text{dist}_{X \times G} \bigcup_{i \in [n]} S^{-i} (P \vee c), \text{dist}_{\bar{X} \times G} \bigcup_{i \in [n]} \bar{S}_k (\bar{P}_k \vee \bar{\beta}_k c) \right\|_{i, \mathcal{M}} = 0,
\]

and

\[
\bar{A}_k \subset \bigcup_{\varepsilon_k \in [-m_k, m_k]} \bar{T}_{k+1}^{-i}(\bar{P}_{k+1}).
\]

Moreover, the \(\varepsilon_k\) are chosen (by reducing all the \(\{\varepsilon_i\}_{i=0}^\infty\) at stage \(k\)) to guarantee that the partitions \(\bar{P}_k\) converge to a partition \(\bar{P}\), the \(\bar{S}_k\) converge to \(\bar{S}\), the functions \(\bar{\beta}_k\) converge to \(\bar{\alpha}\) and so that \(|\bar{P} - \bar{P}| < \varepsilon\),

\[
\int_X \rho (\bar{\alpha} (\bar{x}), id_G) d\bar{\mu} < \varepsilon,
\]

\[
\bar{\mu} \times \lambda \{(\bar{x}, g) \in \bar{X} \times G \mid \bar{S}(\bar{x}, g) \neq \bar{S}_0(\bar{x}, g)\} < \varepsilon
\]

and for each \(k\)

\[
\bar{A}_k \subset 2\varepsilon_k \bigcup_{\varepsilon_k \in [-m_k, m_k]} \bar{T}^{-i}(\bar{P}).
\]

It follows that for all \(n\)

\[
\left\| \text{dist}_{X \times G} \bigcup_{i \in [n]} S^{-i} (P \vee c), \text{dist}_{\bar{X} \times G} \bigcup_{i \in [n]} \bar{S}^{-i} (\bar{P} \vee \bar{\beta} c) \right\|_{i, \mathcal{M}} = 0
\]

and that \(\bar{P}\) is a generator for \(\hat{T}\). From this we conclude that the \(G\)-extension \((\bar{S}, \hat{T}, \bar{\sigma}, \bar{X})\) is \(G\)-isomorphic to \((S, T, \sigma, X)\).

Finally, we use theorem \([3]\) to prove theorem \([2]\)

**Proof.** (of theorem \([2]\)) Let \((S, T, \sigma, X)\) and \((\bar{S}_0, \bar{T}_0, \bar{\sigma}_0, \bar{X})\) be ergodic \(G\)-extensions on \(X \times G\) and \(\bar{X} \times G\), respectively. Fix \(\varepsilon > 0\). Choose \(\delta\) and \(n\) with respect to
$\varepsilon$ as theorem 3. Let $P$ be a finite generator of $T$. Fix $\zeta > 0$ and $N \in \mathbb{N}$ and let $(x, g) \in (X \times G)$ satisfy

$$\left\| \text{dist}_{S[N]}(x, g) \bigvee_{i \in [n]} S^{-i}(P \vee c), \text{dist}_{X \times G}(S^{-i}(P \vee c)) \right\|_\mathcal{M} < \zeta$$

Let $\bar{\tau}$ be a Rokhlin tower of height $N$ for $\bar{S}$, measurable with respect to $\bar{X}$, and define $\bar{\alpha} : \bar{X} \to G$ and $\bar{P}$ so that for each $\bar{x}$ in the base of $\bar{\tau}$, and for all $i \in [0, N - 1]$,

$$(P \vee \bar{\alpha}c) \left( S^i (\bar{x}, id_G) \right) = (P \vee c) \left( S^i (x, g) \right).$$

If $\zeta$ is chosen sufficiently small, and $N$ is sufficiently large, then we obtain condition 5.1 in the hypotheses of theorem 3. (Note that $\text{dist}_{X \times G}(S^{-i}(P \vee c))$ is invariant under right multiplication in the group component, so the use of the single orbit to define $\bar{P}$ and $\bar{\alpha}$ gives the right distribution of $n$—names on $\bar{X} \times G$). The conclusion of theorem 2 follows from the application of theorem 3.

\[\Box\]

References

[AOW] P. Arnoux, D. S. Ornstein, B. Weiss, Cutting and stacking, interval exchanges and geometric models, *Isr. J. Math.*, 50, no. 1-2, (1985), 160-168.

[D1] R. M. Dudley, Distances of probability measures and random variables, *Ann. Math. Statistics*, 39, (1968), 1563-1572.

[D2] R. M Dudley, *Real Analysis and Probability*, Wadsworth &Brooks/Cole, Pacific Grove, CA, 1989.

[F] A. Fieldsteel, Factor orbit equivalence of compact group extensions, *Isr. J. Math.*, 38, no. 4, (1981), 289-303.

[G] M. Gerber, Factor orbit equivalence of compact group extensions and classification of finite extensions of ergodic automorphisms, *Isr. J. Math.*, 57, no. 1, (1987), 28-48.

[R] D. J. Rudolph, Restricted orbit equivalence, *Mem. AMS*, 323 (1985).

[KR1] J. Kammeyer, D.J. Rudolph, Restricted orbit equivalence for ergodic $\mathbb{Z}^d$ actions, I., *Ergodic Th. Dyn. Sys.*, 17, no. 5, 1997, 1083–1129.

[KR2] J. Kammeyer, D.J. Rudolph, *Restricted orbit equivalence for actions of discrete amenable groups*, Cambridge Tracts in Mathematics, 146. Cambridge University Press, 2002.

[O] D. S. Ornstein, *Ergodic Theory, Randomness and Dynamical Systems*, Yale University Press, New Haven, 1970.

Department of Mathematics, California State University, Fresno, CA 93740

E-mail address: ababichev@csufresno.edu

Department of Mathematics and Computer Science, Wesleyan University, Middletown, CT 06459

E-mail address: afieldsteel@wesleyan.edu