Abstract. We construct entropy increasing monotone factors in the context of a Bernoulli shift over the free group of rank at least two.

1. Introduction

Let $\kappa$ be a probability measure on a finite set $K$. We will mainly be concerned with the simple case where $K = \{0, 1\}$, where we call $\kappa(1) := \kappa(\{1\}) \in (0, 1)$ the intensity of $\kappa$. Let $G$ be a group. A Bernoulli shift over $G$ with base $(K, \kappa)$ is the measure-preserving system $(G, K^G, \kappa^G)$, where $G$ acts on $K^G$ via $(gx)(f) = x(g^{-1}f)$ for $x \in K^G$ and $g, f \in G$. Let $\iota$ be a probability measure of lower intensity. We say that a measurable map $\phi : K^G \to K^G$ is an equivariant thinning from $\kappa$ to $\iota$ if $\phi(x)(g) \leq x(g)$ for all $x \in K^G$ and $g \in G$, the push-forward of $\kappa^G$ under $\phi$ is $\iota^G$, and $\phi$ is equivariant $\kappa^G$-almost-surely; that is, on a set of full-measure, $\phi \circ g = g \circ \phi$ for all $g \in G$.

**Theorem 1.** Let $\kappa$ and $\iota$ be probability measures on $\{0, 1\}$ and $\iota$ be of lower intensity. For Bernoulli shifts over the free group of rank at least two, there exists an equivariant thinning from $\kappa$ to $\iota$.

Theorem 1 does not hold with such generality in the case of a Bernoulli shift over an amenable group like the integers. Recall that the entropy of a probability measure $\kappa$ on a finite set $K$ is given by

$$H(\kappa) := -\sum_{i \in K} \kappa(i) \log \kappa(i).$$

**Theorem 2** (Ball [3], Soo [16]). Let $\kappa$ and $\iota$ be probability measures on $\{0, 1\}$ and $\iota$ be of lower intensity. For Bernoulli shifts over the integers, there exists an equivariant thinning from $\kappa$ to $\iota$ if and only if $H(\kappa) \geq H(\iota)$.

2010 *Mathematics Subject Classification.* 37A35, 60G10, 60E15.

*Key words and phrases.* Sinai factor theorem, Ornstein and Weiss, thinning, stochastic domination, monotone coupling, non-amenable group.

Funded in part by a New Faculty General Research Fund.
In Theorem 2, the necessity of $H(\kappa) \geq H(\iota)$ follows easily from the classical theory of Kolmogorov-Sinai entropy [8, 19], which we now recall. Let $G$ be a group and let $\kappa$ and $\iota$ be probability measures on a finite set $K$. An equivariant map $\phi$ is a factor from $\kappa$ to $\iota$ if the push-forward of $\kappa^G$ under $\phi$ is $\iota^G$, and is an isomorphism if $\phi$ is a bijection and its inverse also serves as a factor from $\iota$ to $\kappa$. In the case $G = \mathbb{Z}$, Kolmogorov proved that entropy is non-increasing under factor maps; this implies the necessity of $H(\kappa) \geq H(\iota)$ in Theorem 2. Furthermore, Sinai [15] proved that there is a factor from $\kappa$ to $\iota$ if $H(\kappa) \geq H(\iota)$, and Ornstein [12] proved there is an isomorphism from $\kappa$ to $\iota$ if and only if $H(\kappa) = H(\iota)$. Thus entropy is a complete invariant for Bernoulli shifts over $\mathbb{Z}$. Ornstein and Weiss [13] generalized these results to the case where $G$ is an amenable group. See also Keane and Smorodinsky for concrete constructions of factor maps and isomorphisms [9, 10].

The sufficiency of $H(\kappa) > H(\iota)$ in Theorem 2 was first proved by Ball [3]. The existence of an isomorphism that is also an equivariant thinning in the equal entropy case was proved by Soo [16]. Let us remark that the factor maps given in standard proofs of the Sinai and Ornstein theorems will not in general be monotone; that is, they may not satisfy $\phi(x)(i) \leq x(i)$ for all $x \in \{0, 1\}^\mathbb{Z}$ and $i \in \mathbb{Z}$.

Towards the end of their 1987 paper, Ornstein and Weiss [13] give a simple but remarkable example of an entropy increasing factor in the case where $G$ is the free group of rank at least two, which is further elaborated upon by Ball [2]. It was an open question until recently whether all Bernoulli shifts over a free group of rank at least two are isomorphic. This question was answered negatively by Lewis Bowen [5] in 2010, who proved that although entropy can increase under factor maps, in the context of a free group with rank at least two, it is still a complete isomorphism invariant. Recently, there has been much interest in studying factors in the non-amenable setting; see Russell Lyons [11] for more information.

Our proof of Theorem 1 will make use of a variation of the Ornstein and Weiss example in Ball [2] and a primitive version of a marker-filler type construction, in the sense of Keane and Smorodinsky [9, 10]. Our construction uses randomness already present in the process in a careful way as to mimic a construction that one would make if additional independent randomization were available. This approach was taken by Holroyd, Lyons, and Soo [7], Angel, Holroyd, and Soo [1], and Ball [4] for defining equivariant thinning in the context of Poisson point processes.
2. Tools

2.1. Coupling. Let \((A, \alpha)\) and \((B, \beta)\) be probability spaces. A **coupling** of \(\alpha\) and \(\beta\) is a probability measure on the product space \(A \times B\) which has \(\alpha\) and \(\beta\) as its marginals. For a random variable \(X\), we will refer to the measure \(\mathbb{P}(X \in \cdot)\) as the **law** or the **distribution** of \(X\).

If two random variables \(X\) and \(Y\) have the same law, we write \(X \overset{d}{=} Y\). Similarly, a **coupling** of random variables \(X\) and \(Y\) is a pair of random variables \((X', Y')\), where \(X'\) and \(Y'\) are defined on the same probability space and have the same law as \(X\) and \(Y\), respectively. Thus a coupling of random variables gives a coupling of the laws of the random variables. Often we will refer to the law of a pair of random variables as the **joint distribution** of the random variables. In the case that \(A = B\) and \(A\) is a partially ordered by the relation \(\preceq\), we say that a coupling \(\gamma\) is **monotone** if \(\gamma\{(a, b) \subset A \times A : b \preceq a\} = 1\). We will always endow the space of binary sequences \(\{0, 1\}^I\) indexed by a set \(I\) with the partial order \(x \preceq y\) if and only if \(x_i \leq y_i\) for \(i \in I\).

**Example 3** (Independent thinning). Let \(\kappa\) and \(\iota\) be probability measures on \(\{0, 1\}\), where \(\kappa(1) := p \geq \iota(1) := q\). Let \(r := \frac{p - q}{p}\). Then the measure \(\rho\) on \(\{0, 1\}^2\) given by

\[
\rho(0, 0) = 1 - p, \quad \rho(0, 1) = 0, \quad \rho(1, 0) = rp, \quad \text{and} \quad \rho(1, 1) = (1 - r)p
\]

is a monotone coupling of \(\kappa\) and \(\iota\). Thus under \(\rho\), a 1 is thinned to a 0 with probability \(r\) and kept with probability \(1 - r\). Clearly, the product measure \(\rho^n\) is a monotone coupling of \(\kappa^n\) and \(\iota^n\). We will refer to the coupling \(\rho^n\) as the **independent thinning of \(\kappa^n\) to \(\iota^n\)**.

The following simple lemma is one of the main ingredients in the proof of Theorem \(\Box\). In it we construct a coupling of \(\kappa^n\) and \(\iota^n\) for \(n\) sufficiently large which will allow us to extract spare randomness from a related coupling of \(\kappa^G\) and \(\iota^G\). We will write \(0^n1^m\) to indicate the binary sequence of length \(n + m\) of \(n\) zeros followed by \(m\) ones.

**Lemma 4** (Key coupling). Let \(\kappa\) and \(\iota\) be probability measures on \(\{0, 1\}\), where \(\kappa\) is of greater intensity. For \(n\) sufficiently large, there exists a monotone coupling \(\gamma\) of \(\kappa^n\) and \(\iota^n\) such that

\[
\gamma(10^n{-2}, 0^n) = \kappa^n(10^n{-2})
\]

and

\[
\gamma(010^n{-2}, 0^n) = \kappa^n(010^n{-2}).
\]

**Proof.** Let \(p = \kappa(1)\), \(q = \iota(1)\), and \(\rho^n\) be the independent thinning of \(\kappa^n\) to \(\iota^n\) as in Example \(\Box\). We will perturb \(\rho^n\) to give the required coupling.
We specify a probability measure \( \varrho \) on \( \{0, 1\}^n \times \{0, 1\}^n \) by stating that it agrees with \( \rho^n \) except on the points \((100^{n-2}, 0^n), (010^{n-2}, 0^n), (100^{n-2}, 100^{n-2}), \) and \((010^{n-2}, 010^{n-2})\), where we specify that

\[
\varrho(100^{n-2}, 0^n) = \varrho(010^{n-2}, 0^n) = p(1 - p)^{n-1}
\]

and

\[
\varrho(100^{n-2}, 100^{n-2}) = \varrho(010^{n-2}, 010^{n-2}) = 0.
\]

Thus \( \varrho \) is almost a monotone coupling of \( \kappa^n \) and \( \iota^n \), except that from our changes to \( \rho^n \) we have

\[
\sum_{x \in \{0, 1\}^n} \varrho(x, 0^n) = \sum_{x \in \{0, 1\}^n} \rho^n(x, 0^n) - \rho^n(100^{n-2}, 0^n) - \rho^n(010^{n-2}, 0^n)
\]

\[
\quad + \varrho(100^{n-2}, 0^n) + \varrho(010^{n-2}, 0^n)
\]

\[
= (1 - q)^n + 2p(1 - p)^{n-1}(1 - r),
\]

and

\[
\sum_{x \in \{0, 1\}^n} \varrho(x, 100^{n-2}) = \sum_{x \in \{0, 1\}^n} \rho^n(x, 100^{n-2}) - \rho^n(100^{n-2}, 100^{n-2})
\]

\[
\quad + \varrho(100^{n-2}, 100^{n-2})
\]

\[
= q(1 - q)^{n-1} - p(1 - p)^{n-1}(1 - r) + 0
\]

\[
= \sum_{x \in \{0, 1\}^n} \varrho(x, 010^{n-2}),
\]

where \( r = \frac{p - q}{p} \).

We perturb \( \varrho \) to obtain the desired coupling \( \gamma \). Consider the set \( B_1 \) of all binary sequences of length \( n \), where \( x \in B_1 \) if and only if \( x_1 = 1 \), \( x_2 = 0 \), and \( \sum_{i=3}^{n} x_i = 1 \). Similarly, let \( B_2 \) be the set of all binary sequences of length \( n \), where \( x \in B_2 \) if and only if \( x_1 = 0 \), \( x_2 = 1 \), and \( \sum_{i=3}^{n} x_i = 1 \). The sets \( B_1 \) and \( B_2 \) are disjoint, and each have cardinality \( n - 2 \).

For \( x \in B_1 \cup B_2 \),

\[
\varrho(x, 0^n) = \rho^n(x, 0^n) = p^2(1 - p)^{n-2}r^2,
\]

for \( x \in B_1 \),

\[
\varrho(x, 100^{n-2}) = \rho^n(x, 100^{n-2}) = p^2(1 - p)^{n-2}r(1 - r),
\]

and for \( x \in B_2 \),

\[
\varrho(x, 010^{n-2}) = \rho^n(x, 010^{n-2}) = p^2(1 - p)^{n-2}r(1 - r).
\]

Note that for \( n \) sufficiently large

\[
\sum_{x \in B_1 \cup B_2} \varrho(x, 0^n) = 2(n - 2)p^2(1 - p)^{n-2}r^2 > 2p(1 - p)^{n-1}(1 - r).
\]
Let $γ$ be equal to $ρ$ except on the set of points
\[ \{ (x, 0^n) : x \in B_1 \cup B_2 \} \cup \{ (x, 100^{n-2}) : x \in B_1 \} \cup \{ (x, 010^{n-2}) : x \in B_2 \}, \]
where we make the following adjustments. For $x \in B_1 \cup B_2$, set
\[
γ(x, 0^n) = p^2(1 - p)^{n-2}r^2 - \frac{p(1-p)^{n-1}(1-r)}{n-2} > 0,
\]
for $x \in B_1$, set
\[
γ(x, 100^{n-2}) = p^2(1 - p)^{n-2}(1-r)r + \frac{p(1-p)^{n-1}(1-r)}{n-2},
\]
and for $x \in B_2$, set
\[
γ(x, 010^{n-2}) = p^2(1 - p)^{n-2}(1-r)r + \frac{p(1-p)^{n-1}(1-r)}{n-2}.
\]
That $γ$ has the required properties follows from its construction. \(\square\)

To illustrate the utility of Lemma 4, we will give a different proof of the following result of Peled and Gurel-Gurevich [6]. Let $\mathbb{N} = \{0, 1, 2, \ldots\}$.

**Theorem 5** (Peled and Gurel-Gurevich [6]). Let $κ$ and $ι$ be probability measures on $\{0, 1\}$, where $κ$ is of greater intensity. There exists a measurable map $φ : \{0, 1\}^{\mathbb{N}} \to \{0, 1\}^{\mathbb{N}}$ such that the push-forward of $κ^{\mathbb{N}}$ under $φ$ is $ι^{\mathbb{N}}$ and $φ(x)(i) ≤ x(i)$ for all $x \in \{0, 1\}^{\mathbb{N}}$ and all $i \in \mathbb{N}$.

We note that in [6, Theorem 1.3], they use the dual terminology of thickenings; their equivalent theorem states that for probability measures $ι$ and $κ$ on $\{0, 1\}$, where $ι$ is of lesser intensity, there is a measurable map $φ : \{0, 1\}^{\mathbb{N}} \to \{0, 1\}^{\mathbb{N}}$ such that the push-forward of $ι^{\mathbb{N}}$ under $φ$ is $κ^{\mathbb{N}}$ and $φ(x)(i) ≥ x(i)$ for all $x \in \{0, 1\}^{\mathbb{N}}$ and all $i \in \mathbb{N}$.

In the proof of Theorem 5, we will make use of the following two lemmas. We say that a random variable $U$ is **uniformly distributed** in $[0, 1]$ if the probability that $U$ lies in a Borel subset of the unit interval is given by the Lebesgue measure of the set.

**Lemma 6.** Let $(X, Y)$ be a pair of discrete random variables taking values on the finite set $A \times B$ with joint distribution $γ$. There exists a measurable function $Γ : A \times [0, 1] \to B$ such that if $U$ is uniformly distributed in $[0, 1]$ and independent of $X$, then $(X, Γ(X, U))$ has joint distribution $γ$.

**Proof.** Assume that $\mathbb{P}(X = a) > 0$, for all $a \in A$. Let $B = \{b_1, \ldots, b_n\}$. For each $a \in A$, let
\[
q_a(j) := \mathbb{P}(Y \in \{b_1, \ldots, b_j\} | X = a) = \frac{\mathbb{P}(Y \in \{b_1, \ldots, b_j\}, X = a)}{\mathbb{P}(X = a)}.
\]
for all $1 \leq j \leq n$. Set $q_a(0) = 0$ and note that $q_a(n) = 1$, so that
\[
\mathbb{P}(q_a(j - 1) \leq U < q_a(j)) = \frac{\mathbb{P}(Y = b_j, X = a)}{\mathbb{P}(X = a)}.
\]
For each $1 \leq j \leq n$, let
\[
\Gamma(a, u) := b_j \text{ if } q_a(j - 1) \leq u < q_a(j).
\]
We call a $\{0, 1\}$-valued random variable a Bernoulli random variable. The following lemma allows us to code sequences of independent coin-flips into sequences of uniformly distributed random variables.

**Lemma 7.** There exists a measurable function $c : \{0, 1\}^N \rightarrow [0, 1]^N$ such that if $B = (B_i)_{i \in \mathbb{N}}$ is a sequence of i.i.d. Bernoulli random variables with mean $\frac{1}{2}$, then $(c(B_i))_{i \in \mathbb{N}}$ is a sequence of i.i.d. random variables that are uniformly distributed in $[0, 1]$.

**Proof.** The result follows from the Borel isomorphism theorem. See [17, Theorem 3.4.23] for more details. □

**Proof of Theorem 5.** Let $\gamma$ be the monotone coupling of $\kappa^n$ and $\iota^n$ given by Lemma 4, so that $\gamma$ is a measure on $\{0, 1\}^n \times \{0, 1\}^n \equiv (\{0, 1\} \times \{0, 1\})^n$. Thus the product measure $\gamma^2$ is a monotone coupling of $\kappa^{2n}$ and $\iota^{2n}$ and $\gamma^N$ gives a monotone coupling of $\kappa^N$ and $\iota^N$. We will modify the coupling $\gamma^N$ to become the required map $\phi$. In order to do this, it will be easier to think in terms of random variables rather than measures.

Let $X = (X_i)_{i \in \mathbb{N}}$ be an i.i.d. sequence of Bernoulli random variables with mean $\kappa(1)$. For each $j \geq 0$, let
\[
X^j := (X_{jn}, \ldots, X_{(j+1)n-1}),
\]
so that the random variables are partitioned into blocks of size $n$. Let $U = (U_i)_{i \in \mathbb{N}}$ be an i.i.d. sequence of random variables that are uniformly distributed in $[0, 1]$. Also assume that $U$ is independent of $X$, and let $Y = (Y_i)_{i \in \mathbb{N}}$ be an i.i.d. sequence of Bernoulli random variables with mean $\iota(1)$.

By Lemmas 4 and 6 let $\Gamma : \{0, 1\}^n \times [0, 1] \rightarrow \{0, 1\}^n$ be a measurable map such that $(X^1, \Gamma(X^1, U_1))$ has joint law $\gamma$ and $\Gamma(w, v) = 0^n$ for all $v \in [0, 1]$ if $w \in \{100^{n-2}, 010^{n-2}\}$. We have that
\[
(X, (\Gamma(X^i, U_i))_{i \in \mathbb{N}})
\]
gives a monotone coupling of $X$ and $Y$ with law $\gamma^N$.

For each $j \in \mathbb{N}$, call $X^j$ **special** if $X^j \in \{100^{n-2}, 010^{n-2}\}$ and let $S \subset \mathbb{N}$ be the random set of $j \in \mathbb{N}$ for which $X^j$ are special. Note that almost surely, $S$ is an infinite set. Let $\bar{X} = (\bar{X}_i)_{i \in \mathbb{N}}$ be the sequence of
binary digits such that \( X^j = X^j \) if \( j \not\in S \) and \( X^j = 0^n \) if \( j \in S \). We have that
\[
(\Gamma(X^i, U_i))_{i \in \mathbb{N}} = (\Gamma(\tilde{X}^i, U_i))_{i \in \mathbb{N}}.
\]

Let \((s_i)_{i \in \mathbb{N}}\) be the enumeration of \( S \), where \( s_0 < s_1 < s_2 < s_3 \ldots \). Consider the sequence of random variables given by
\[
b(X) := (1[X^s_i = 100^{n-2}])_{i \in \mathbb{N}} = (X_{s,n})_{i \in \mathbb{N}}
\]
Since \(100^{n-2} \) and \( 010^{n-2} \) occur with equal probability, we have that \( b(X) \) is an i.i.d. sequence of Bernoulli random variables with mean \( \frac{1}{2} \). Furthermore, we have that \( b(X) \) is independent of \( \tilde{X} \), since \( b(X) \) only depends on the values of \( X \) on the special blocks. Let \( c \) be the function from Lemma \( 7 \) so that \( c(b(X)) \equiv d U \). Since \( b(X) \) is independent of \( \tilde{X} \),
\[
[\Gamma(X^i, U_i)]_{i \in \mathbb{N}} = [\Gamma(\tilde{X}^i, U_i)]_{i \in \mathbb{N}} \\
\equiv [\Gamma(\tilde{X}^i, c(b(X)))_{i \in \mathbb{N}} \\
= [\Gamma(X^i, c(b(X)))_{i \in \mathbb{N}}.
\]
Thus \( \left( X, [\Gamma(X^i, c(b(X)))_{i \in \mathbb{N}} \right) \) is another monotone coupling of \( X \) and \( Y \). Hence, we define
\[
\phi(x) := [\Gamma(x^i, c(b(x)))_{i \in \mathbb{N}}
\]
for all \( x \in \{0, 1\}^\mathbb{N} \) when the set \( S \) is infinite, and set \( \phi(x) = 0^\mathbb{N} \) when \( S \) is finite—an event that occurs with probability zero. \( \square \)

2.2. Joinings. Let \( T : \{0, 1\}^\mathbb{Z} \to \{0, 1\}^\mathbb{Z} \) be the left-shift given by \( (Tx)_i = x_{i+1} \) for all \( x \in \{0, 1\}^\mathbb{Z} \) and all \( i \in \mathbb{Z} \). Let \( \kappa \) and \( \iota \) be probability measures on \( \{0, 1\} \). A joining of \( \kappa^\mathbb{Z} \) and \( \iota^\mathbb{Z} \) is a coupling \( \varphi \) of the two measures with the additional property that \( \varphi \circ (T \times T) = \varphi \). We will make use of the following joining in the proof of Theorem \( 1 \).

Example 8. Let \( \kappa \) and \( \iota \) be probability measures on \( \{0, 1\} \). Assume that the intensity of \( \kappa \) is greater than the intensity of \( \iota \). Let \( x \in \{0, 1\}^\mathbb{Z} \), and let \( n \) be sufficiently large as in Lemma \( 4 \). Call the subset \([j, j + 2n + 1] \subset \mathbb{Z} \) a marker if \( x_i = 0 \) for all \( i \in [j, j + 2n] \) and \( x_{j+2n+1} = 1 \). Notice that two distinct markers have an empty intersection. Call an interval a filler if it is nonempty and lies between two markers. Thus each \( x \in \{0, 1\}^\mathbb{Z} \) partitions \( \mathbb{Z} \) into intervals of markers and fillers. Call a filler fitted if it is of size \( n \), and call a filler special if it is both fitted and of the form \( 100^{n-2} \) or \( 010^{n-2} \).

Let \( X \) have law \( \kappa^\mathbb{Z} \) and \( Y \) have law \( \iota^\mathbb{Z} \). In what follows we describe explicitly how to obtain a monotone joining of \( X \) and \( Y \), where the independent thinning is used everywhere, except at the fitted fillers,
where the coupling from Lemma 4 is used. Let \( U = (U_i)_{i \in \mathbb{Z}} \) be an i.i.d. sequence of random variables that are uniformly distributed in \([0, 1]\) and independent of \( X \). By Example 3 and Lemma 6 let \( R : \{0, 1\} \times [0, 1] \to \{0, 1\} \) be a measurable function such that \( R(X_i, U_i) \leq X_i \) is a Bernoulli random variable with mean \( \nu(1) \). Let \( \Gamma \) and \( \gamma \) be as in the proof of Theorem 5, so that \((X_1, \ldots, X_n, \Gamma(X_1, \ldots, X_n, U_1))\) has law \( \gamma \). Consider the function \( \Phi : \{0, 1\}^\mathbb{Z} \times [0, 1]^\mathbb{Z} \to \{0, 1\}^\mathbb{Z} \) defined by \( \Phi(x, u) = R(x_i, u_i) \) if \( i \) is not in a fitted filler. For \((j, j+1, \ldots, j+n)\) in a fitted filler, we set \( \Phi(x, u)_j, \ldots, \Phi(x, u)_{j+n} = \Gamma(x_j, \ldots, x_{j+n}, u_j) \).

The law of \( X \) restricted to a filler interval is the law of a finite sequence of i.i.d. Bernoulli random variables with mean \( \kappa(1) \), conditioned not to contain a marker. Note that since a fitted interval is of size \( n \), and a marker is of size \( 2n + 1 \), the law of \( X \) restricted to a fitted interval is just the law of a finite sequence of i.i.d. Bernoulli random variables with mean \( \kappa(1) \). Furthermore, conditioned on the locations of the markers, the restrictions of \( X \) to each filler interval are independent (see for example Keane and Smorodinsky [9, Lemma 4] for a detailed proof).

Hence, \( \Phi(X, U) \overset{d}{=} Y \). In addition, since all the couplings involved are monotone, we easily have that \( \Phi(X, U)_i \leq X_i \) for all \( i \in \mathbb{Z} \).

Remark 9. To emphasize the strong form of independence in Example 8, we note that if \( A = (A_i)_{i \in \mathbb{Z}} \) are independent Bernoulli random variables with mean \( \frac{1}{2} \) that are independent of \( X \), then \( (A_{jn})_{j \in S} \) has the same law as \( (X_{jn})_{j \in S} \). Recall if \( j \in S \) then \( X^j = (X_{jn}, \ldots, X_{(j+1)n-1}) \) is special. In addition, if \( X' \) is such that \( X'_i = X_i \) for every \( i \) not in a special filler of \( X \) and on each special filler of \( X \) we set \( X'_{jn} = A_{jn} \), \( X'_{jn+1} = 1 - A_{jn} \), and

\[
X'_{jn+2} = X'_{jn+3} = \cdots = X'_{(j+1)n-1} = 0,
\]

then \( X' \overset{d}{=} X \). Thus we can independently resample on the special fillers without affecting the distribution of \( X \). ⊓⊔

2.3. The example of Ornstein and Weiss. Let \( \mathbb{F}_r \), be the free group of rank \( r \geq 2 \). Let \( a \) and \( b \) be two of its generators. The Ornstein and Weiss [13] entropy increasing factor map is given by

\[
\phi(x)(g) = (x(g) \oplus x(ga), x(g) \oplus x(gb))
\]

for all \( x \in \{0, 1\}^{\mathbb{F}_r} \) and all \( g \in \mathbb{F}_2 \), where

\[
\phi : \{0, 1\}^{\mathbb{F}_r} \to (\{0, 1\} \times \{0, 1\})^{\mathbb{F}_r} \equiv \{00, 01, 10, 11\}^{\mathbb{F}_r}.
\]
pushes the uniform product measure \( \left( \frac{1}{2}, \frac{1}{2} \right)^{\mathbb{F}_r} \) forward to the uniform
product measure \( \left( \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4} \right)^{\mathbb{F}_r} \); the required independence follows from
the observation that if \( m \oplus n := m + n \mod 2 \), if \( X, X' \), and \( Y \) are
independent Bernoulli random variables with mean \( \frac{1}{2} \), and if \( Z := X \oplus Y \) and \( Z' := X' \oplus Y \), then \( Z \) and \( Z' \) are independent, even though
they both depend on \( Y \).

Ornstein and Weiss’s example can be iterated to produce an infinite
number of bits at each vertex in the following way. As in Ball [2,
Proposition 2.1], we will define \( \phi_k : \{0,1\}^{\mathbb{F}_r} \to \{0,1\}^k \mathbb{F}_r \) inductively
for \( k \geq 2 \). Let \( \tilde{\phi}_k : \{0,1\}^{\mathbb{F}_r} \to \{0,1\}^{\mathbb{F}_r} \) be the last coordinate of \( \phi_k \)
so that \( \tilde{\phi}_k(x)(g) = [\phi_k(x)(g)]_k \) for all \( x \in \{0,1\}^{\mathbb{F}_r} \) and all \( g \in \mathbb{F}_2 \). Set
\( \phi_2 = \phi \). For \( k \geq 3 \), let \( \phi_k \) be given by
\[
\phi_k(x)(g) = \left([\phi_{k-1}(x)(g)]_1, \ldots, [\phi_{k-1}(x)(g)]_{k-2}, (\phi \circ \tilde{\phi}_{k-1})(x)(g)\right)
\]
for all \( x \in \{0,1\}^{\mathbb{F}_r} \) and all \( g \in \mathbb{F}_2 \). At each step we are saving one bit
to generate two new bits using the original map \( \phi \). The map \( \phi_k \) pushes
the uniform product measure \( \left( \frac{1}{4}, \frac{1}{4} \right)^{\mathbb{F}_r} \) forward to the uniform product
measure on \( \{0,1\}^k \mathbb{F}_r \). By taking the limit, we obtain the mapping
\[
\phi_\infty : \{0,1\}^{\mathbb{F}_r} \to \{0,1\}^{\mathbb{Z}^+} \mathbb{F}_r
\]
which yields a sequence of i.i.d. fair bits at each coordinate \( g \in \mathbb{F}_2 \),
independently. Note that \( \phi_\infty(x)(g)_k = \phi_n(x)(g)_k \) for all \( n > k \). In our
proof of Theorem 1 we will use this iteration, which Ball attributes to
Timár.

3. PROOF OF THE MAIN THEOREM

Proof of Theorem 1 Let \( r \geq 2 \). We begin by extending the same mono-
tone joining defined in Example 8 to a monotone joining of \( \kappa \mathbb{F}_r \) and \( \nu \mathbb{F}_r \).
Let \( X \) have law \( \kappa \mathbb{F}_r \) and \( Y \) have law \( \nu \mathbb{F}_r \); then \( X = (X_g)_{g \in \mathbb{F}_r} = (X(g))_{g \in \mathbb{F}_r} \) are i.i.d. Bernoulli random variables with mean \( \kappa(1) \). As in
the Ornstein and Weiss example, it will be sufficient to use only two
generators \( a \) and \( b \) in the expression of our equivariant thinning. We
refer to the string of generators and their inverses that make up the
representation of an element in \( \mathbb{F}_r \) as a word, and the individual gen-
erators and inverses as letters. We call a word reduced if its string
of letters has no possible cancellations.

Consider \( \mathbb{F}_r \) as being partitioned into infinitely many \( \mathbb{Z} \) copies \( Z(w) \)
in the following way. Let \( \mathbb{F}_r' \) be the set of reduced words in \( \mathbb{F}_r \), that
do not end in either \( b \) or \( b^{-1} \). For each \( w \in \mathbb{F}_r' \), set \( Z(w) := \{wb^i\}_{i \in \mathbb{Z}} \).
Indeed, any element in \( \mathbb{F}_r' \) may be written as \( wb^i \) for unique reduced
\( w \in \mathbb{F}_r \) and \( i \in \mathbb{Z} \).
Let \( n \) be sufficiently large for the purposes of Lemma 4. We define markers, fillers, fitted fillers, and special fillers on each of the \( \mathbb{Z} \) copies in the obvious way. For example, if \( x \in \{0, 1\}^{F_r} \) and \( w \in \mathbb{F}_r \), then the set \( \{wb^j, \ldots, wb^{j+2n+1}\} \) is a marker if \( x(wb^j) = 0 \) for all \( i \in [j, 2n] \) and \( x(wb^{2n+1}) = 1 \).

Let \( U' = (U'_g)_{g \in \mathbb{F}_r} \) be i.i.d. uniform random variables independent of \( X \). Let \( \Phi \) be as in Example 8. Define \( \hat{\Phi} : \{0, 1\}^{F_r} \times [0, 1]^{F_r} \to \{0, 1\}^{F_r} \) by

\[
\hat{\Phi}(x, u')_{wb} = \Phi(x(Z(w)), u'(Z(w)))_i
\]

for all \( w \in \mathbb{F}_r \) and all \( i \in \mathbb{Z} \), where \( x(Z(w)) := (x(wb^j))_{j \in \mathbb{Z}} \) and \( u'(Z(w)) := (u'(wb^j))_{j \in \mathbb{Z}} \). Thus we have the monotone joining \( \Phi \) on each \( \mathbb{Z} \) copy \( Z(w) \) in \( \mathbb{F}_r \), so that

\[
\hat{\Phi}(X, U') \overset{d}{=} Y
\]

and \( \hat{\Phi}(X, U')_g \leq X_g \) for all \( g \in \mathbb{F}_r \). Additionally, since \( \Phi \) is a joining, the joint law of \( (X, \hat{\Phi}(X, U')) \) is invariant under \( \mathbb{F}_r \)-actions.

Recall that a special filler has length exactly \( n \), and the filler has two choices of values \( 010^{n-2} \) or \( 100^{n-2} \), which occur with equal probability. We define an initial vertex of a special filler in \( Z(w) \) to be an element \( wb^\kappa \in Z(w) \) where the entire special filler takes values sequentially at vertices on the minimal path from \( wb^\kappa \) to \( wb^{\kappa+n} \). For each \( x \in \{0, 1\}^{F_r} \), let \( V = V(x) \) be the set of initial vertices in \( \mathbb{F}_r \). Note that as in Example 8 the law of \( X \) restricted to a fitted interval is just the law of a finite sequence of i.i.d. Bernoulli random variables with mean \( \kappa(1) \). Furthermore, conditioned on the locations of the markers, the restrictions of \( X \) to each filler interval are independent. Thus for all \( v \in V(X) \), \( X(v) \) is a Bernoulli random variable with mean \( \frac{1}{2} \), and conditioned on \( V(X) \), the random variables \( (X(v))_{v \in V} \) are independent.

We have the same strong form of independence here as emphasized in Remark 3 for Example 8 again by Keane and Smorodinsky [9, Lemma 4]. This is key in our construction: we will use the Bernoulli random variables \( (X(v))_{v \in V} \) to build deterministic substitutes for \( U' \).

Now we adapt the iteration of the Ornstein and Weiss example to assign a sequence of i.i.d. Bernoulli random variables to each \( v \in V \). For each \( v \in V \), let \( k \) be the smallest positive integer such that \( v\alpha^k \in V \); set \( \alpha(v) = v\alpha^k \). Similarly, let \( k' \) be the smallest positive integer such that \( v\beta^k \in V \) and set \( \beta(v) = v\beta^k \). For each \( v \in V \), define

\[
\psi(x)(v) = (x(v) \oplus x(\alpha(v)), x(v) \oplus x(\beta(v))).
\]

Conditioned on \( V \), we have that \( \{(\psi(X))_{v \in V} \) is a family of independent random variables uniformly distributed on \( \{00, 01, 10, 11\} \). We iterate
the map \( \psi \) as we did with the Ornstein and Weiss map \( \phi \). Set \( \psi_2 = \psi \). For \( k \geq 3 \), let

\[
\psi_k(x)(v) = \left( [\psi_{k-1}(x)(v)]_1, \ldots, [\psi_{k-1}(x)(v)]_{k-2}, (\psi \circ \tilde{\psi}_{k-1})(x)(v) \right),
\]

where \( \tilde{\psi}_{k-1}(x)(v) = [\psi_{k-1}(x)(v)]_{k-1} \) is the last coordinate of \( \psi_k \). Let \( \psi_\infty \) be the limit, and let \( B_v = \psi_\infty(X)(v) \), so that conditioned on \( V \), the random variables \( (B_v)_{v \in V} \) are independent, and each \( B_v \) is an i.i.d. sequence of Bernoulli random variables with mean \( \frac{1}{2} \).

For all \( x \in \{0,1\}^F \), let \( \bar{x}(g) = x(g) \) for all \( g \) not in a special filler, and let \( \bar{x}(g) = 0 \) if \( g \) belongs to a special filler. It follows from Remark 9 that if \( B' = (B'_g)_{g \in F} \) are independent Bernoulli random variables with mean \( \frac{1}{2} \) independent of \( X \), then \( (B'_v)_{v \in V(X)} \) has the same law as \( (B_v)_{v \in V(X)} \). Moreover,

\[
(\bar{X}, (B_v)_{v \in V(X)}) \overset{d}{=} (\bar{X}, (B'_v)_{v \in V(X)}). \tag{2}
\]

We assign, in an equivariant way, one uniform random variable to each element in \( F \), using the randomness provided by \( (B_v)_{v \in V} \). Let \( c : \{0,1\}^N \rightarrow [0,1]^N \) be the function from Lemma 7 and let \( g \in F \). Then almost surely there exist \( v \in V \) and a minimal \( j > 0 \) such that \( gb_1 = v \); set \( U_g = c(B_v)_j \). Define \( u : \{0,1\}^F \rightarrow [0,1]^F \) by setting \( u(X) := (U_g)_{g \in F} \). Recall that \( U' = (U'_g)_{g \in F} \) are independent random variables uniformly distributed in \([0,1]\) independent of \( X \). From (2),

\[
(\bar{X}, u(X)) \overset{d}{=} (\bar{X}, U'). \tag{3}
\]

Let \( R : \{0,1\} \times [0,1] \rightarrow \{0,1\} \) and \( \Gamma : \{0,1\}^n \times [0,1] \rightarrow \{0,1\}^n \) be the functions that appear in the definition of \( \Phi \) in Example 8. Recall that \( R \) facilitated independent thinning and \( \Gamma \) the key monotone coupling of Lemma 4. Also recall \( \Gamma(100^{n-2}, t) = 0 = \Gamma(010^{n-2}, t) \) for all \( t \in [0,1] \).

Now define \( \phi : \{0,1\}^F_r \rightarrow \{0,1\}^F_r \) by

\[
\phi(x)(g) = R(x(g), u(x)(g))
\]

for \( g \) not in a fitted filler; if \( \{wb^i, \ldots, wb^{i+n-1}\} \) is a fitted filler, then set

\[
(\phi(x)(wb^i), \ldots, \phi(x)(wb^{i+n-1})) = \Gamma(x(wb^i), \ldots, x(wb^{i+n-1}), u(x)(wb^i)).
\]

Note \( \phi \) is defined so that \( \phi(x) = \hat{\Phi}(x, u(x)) \). The map \( \phi \) is equivariant and satisfies \( \phi(x)(g) \leq x(g) \) by construction. It remains to verify that \( \phi(X) \overset{d}{=} Y \).

By the definition of \( \Gamma \), we have \( \phi(X) = \hat{\Phi}(X, u(X)) \); that is, all special fillers are sent to \( 0^n \). A similar remark applies to the map \( \hat{\Phi} \). From (1)
\[ \phi(X) = \hat{\Phi}(X, u(X)) = \hat{\Phi}(\bar{X}, u(X)) = \hat{\Phi}(\bar{X}, U') = \hat{\Phi}(X, U') = Y. \]

4. Generalizations and questions

4.1. Stochastic domination. Let \([N] = \{0, 1, \ldots, N - 1\}\) be endowed with the usual total ordering. Let \(\kappa\) and \(\iota\) be probability measures on \([N]\). We say that \(\kappa\) stochastically dominates \(\iota\) if
\[
\sum_{j=0}^{N-1} \kappa_j \leq \sum_{i=0}^{N-1} \iota_i
\]
for all \(j \in [N]\). An elementary version of Strassen’s theorem [18, Theorem 11] gives that \(\kappa\) stochastically dominates \(\iota\) if and only if there exists a monotone coupling of \(\kappa\) and \(\iota\). Notice that in the case \(N = 2\), we have that \(\kappa\) stochastically dominates \(\iota\) if and only if \(\iota\) is not of higher intensity than \(\kappa\). Thus Theorem 1 gives a positive answer to a special case of the following question.

**Question 1.** Let \(\kappa\) and \(\iota\) be probability measures on \([N]\), where \(\kappa\) stochastically dominates \(\iota\), and \(\kappa\) gives positive measure to at least two elements of \([N]\). Let \(G\) be the free group of rank at least two. Does there exist a measurable equivariant map \(\phi : [N]^G \to [N]^G\) such that the push-forward of \(\kappa^G\) is \(\iota^G\) and \(\phi(x)(g) \leq x(g)\) for all \(x \in [N]^G\) and \(g \in G\)?

In Question 1, we call the map \(\phi\) a monotone factor from \(\kappa\) to \(\iota\). A necessary condition for the existence of a monotone factor from \(\kappa\) to \(\iota\) is that \(\kappa\) stochastically dominates \(\iota\). In the case \(G = \mathbb{Z}\), Ball [3] proved that there exists a monotone factor from \(\kappa\) to \(\iota\) provided that \(\kappa\) stochastically dominates \(\iota\), \(H(\kappa) > H(\iota)\), and \(\iota\) is supported on two symbols; Quas and Soo [14] removed the two symbol condition on \(\iota\).

In the non-amenable case, where \(G\) is a free group of rank at least two, one can hope that Question 1 can be answered positively, without any entropy restriction. However, the analogue of Lemma 4 that was key to the proof of Theorem 1 does not apply in the simple case where \(\kappa = (0, \frac{1}{2}, \frac{1}{2})\) and \(\iota = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})\). In particular, for all \(n \geq 1\), there is no coupling \(\rho\) of \(\kappa^n\) and \(\iota^n\) for which there exists \(x \in \{1, 2\}_n\) and \(y \in \{0, 1, 2\}_n\) such that \(\rho(x, y) = \kappa^n(x) = \frac{1}{2}^n\), since \(\rho(x, y) \leq \iota^n(y) = \frac{1}{3}^n\).

4.2. Automorphism-equivariant factors. The Cayley graph of \(F_n\) is the regular tree \(\mathbb{T}_{2n}\) of degree \(2n\). We note that \(F_n\) is a strict subset of the group of graph automorphisms of \(\mathbb{T}_{2n}\). The map that we constructed in Theorem 1 is not equivariant with respect to the full automorphism group of \(\mathbb{T}_{2n}\). In particular, our definition of a marker is not equivariant with respect to the automorphism which exchanges...
a-edges and b-edges in $\mathbb{T}_{2n}$. However, Ball generalizes the Ornstein and Weiss example to the full automorphism group in [2, Theorem 3.3] by proving that for any $d \geq 3$, there exists a measurable mapping $\phi : \{0, 1\}^{T_d} \rightarrow [0, 1]^{T_d}$ which pushes the uniform product measure on two symbols forward to the product measure of Lebesgue measure on the unit interval, equivariant with respect to the group of automorphisms of $\mathbb{T}_d$. Moreover, she proved the analogous result for any tree with bounded degree, no leaves, and at least three ends.

**Question 2.** Let $T$ be a tree with bounded degree, no leaves, and at least three ends. Let $\kappa$ and $\iota$ be probability measures on $\{0, 1\}$ and $\iota$ be of lower intensity. Does there exists a thinning from $\kappa$ to $\iota$ that is equivariant with respect to the full automorphism group of $T$?

**Acknowledgements**

We thank the referee for carefully reviewing the paper and providing various helpful comments and suggestions.

**References**

[1] O. Angel, A. E. Holroyd, and T. Soo. Deterministic thinning of finite Poisson processes. *Proc. Amer. Math. Soc.*, 139(2):707–720, 2011.
[2] K. Ball. Factors of independent and identically distributed processes with non-amenable group actions. *Ergodic Theory Dynam. Systems*, 25(3):711–730, 2005.
[3] K. Ball. Monotone factors of i.i.d. processes. *Israel J. Math.*, 150:205–227, 2005.
[4] K. Ball. Poisson thinning by monotone factors. *Electron. Comm. Probab.*, 10:60–69 (electronic), 2005.
[5] L. P. Bowen. A measure-conjugacy invariant for free group actions. *Ann. of Math. (2)*, 171(2):1387–1400, 2010.
[6] O. Gurel-Gurevich and R. Peled. Poisson thickening. *Israel J. Math.*, 196(1):215–234, 2013.
[7] A. E. Holroyd, R. Lyons, and T. Soo. Poisson splitting by factors. *Ann. Probab.*, 39(5):1938–1982, 2011.
[8] A. Katok. Fifty years of entropy in dynamics: 1958–2007. *J. Mod. Dyn.*, 1(4):545–596, 2007.
[9] M. Keane and M. Smorodinsky. A class of finitary codes. *Israel J. Math.*, 26:352–371, 1977.
[10] M. Keane and M. Smorodinsky. Bernoulli schemes of the same entropy are finitarily isomorphic. *Ann. of Math. (2)*, 109:397–406, 1979.
[11] R. Lyons. Factors of IID on Trees. *Combin. Probab. Comput.*, 26(2):285–300, 2017.
[12] D. Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Advances in Math.*, 4:337–352, 1970.
[13] D. S. Ornstein and B. Weiss. Entropy and isomorphism theorems for actions of amenable groups. *J. Analyse Math.*, 48:1–141, 1987.
[14] A. Quas and T. Soo. A monotone Sinai theorem. *Ann. Probab.*, 44(1):107–130, 2016.
[15] Y. G. Sinai. *Selecta. Volume I. Ergodic theory and dynamical systems*. Springer, New York, 2010.
[16] T. Soo. A monotone isomorphism theorem. *Probab. Theory Related Fields*, 167(3-4):1117–1136, 2017.
[17] S. M. Srivastava. *A Course on Borel Sets*, volume 180 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.
[18] V. Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36:423–439, 1965.
[19] B. Weiss. The isomorphism problem in ergodic theory. *Bull. Amer. Math. Soc.*, 78:668–684, 1972.

Department of Mathematics, University of Kansas, 405 Snow Hall, 1460 Jayhawk Blvd, Lawrence, Kansas 66045-7594

E-mail address: tsoo@ku.edu
URL: http://www.tsoo.faculty.ku.edu
E-mail address: awilkens@ku.edu
URL: http://www.people.ku.edu/~a548w850/