ISRAEL JOURNAL OF MATHEMATICS 150 (2005), 205–227

MONOTONE FACTORS OF I.I.D. PROCESSES

BY

KAREN BALL

Department of Mathematics, Rawles Hall
Indiana University, Bloomington, IN 47405, USA
e-mail: kareball@indiana.edu
and
Institute for Mathematics and its Applications, University of Minnesota
400 Lind Hall, 207 Church St. SE, Minneapolis, MN 55455, USA

ABSTRACT

Let $B(p)$ and $B(q)$ be Bernoulli shifts on $\{0, 1, \ldots, d-1\}^{\mathbb{Z}}$. If $h(p) > h(q)$, it is a classical theorem of Sinai that there is a factor map taking $B(p)$ to $B(q)$. If, in addition, p stochastically dominates q, we can ask whether there is such a factor map ϕ which is monotone: $\phi(x)_i \leq x_i$ for each coordinate i of almost every point x. Here we show that there is a monotone finitary code from $B(p)$ to $B(q)$ in the case where $B(q)$ is a shift on two symbols.

1. Introduction

Let $p = (p_0, p_1, \ldots, p_{d-1})$ and $q = (q_0, q_1, \ldots, q_{d-1})$ be probability vectors, which define a probability measure on the set $A = \{0, 1 \ldots, d-1\}$. Let $B(p) = (X, \mathcal{A}, \mu, T)$ and $B(q) = (Y, \mathcal{B}, \nu, T)$ be the Bernoulli shifts defined by p and q, with

$$X = Y = A^{\mathbb{Z}},$$
$$\mathcal{A} = \mathcal{B} = \text{product } \sigma\text{-algebra on } A^{\mathbb{Z}},$$
$$T = \text{left shift on } A^{\mathbb{Z}},$$
$$\mu = p^{\mathbb{Z}}, \quad \nu = q^{\mathbb{Z}}.$$

A **factor map** from $B(p)$ to $B(q)$ is a map ϕ: $X \to Y$ which is

- measure-preserving: $\mu(\phi^{-1}(B)) = \nu(B)$ for all $B \in \mathcal{B}$, and
- commutes with the dynamics: $\phi \circ T = T \circ \phi$ almost surely.

Received June 18, 2004

There is a standard partial ordering on $A^{\mathbb{Z}}$, with $x \geq y$ if $x_i \geq y_i$ for each coordinate $i \in \mathbb{Z}$. We are interested here in the existence of factor maps ϕ from $B(p)$ to $B(q)$ which have the special property that for $x \in X$, $\phi(x) \leq x$ almost surely. Such a map ϕ is called a **monotone factor**. For example, suppose we have two coins. The first one is a fair coin and so turns up a 1 with probability 1/2 and a 0 otherwise. The second coin turns up a 1 with probability 1/4 and a 0 otherwise. A typical two-sided sequence of tosses for the first coin will have more 1's than a typical sequence of tosses for the second coin. Here we show that we can take a typical sequence for the first coin and choose to turn some of the 1's to 0's, in a deterministic and shift-equivariant fashion, so that the result is a typical sequence for the second coin, almost surely.

Of course, some assumptions are necessary for the existence of monotone factors. First, for any factor to exist from $B(p)$ to $B(q)$, much less a monotone one, $B(p)$ must have at least as much entropy as $B(q)$: $h(\mu) \geq h(\nu)$, where for Bernoulli shifts,

$$h(\mu) = h(p) = -\sum_{i=0}^{d-1} p_i \log p_i \quad \text{and} \quad h(\nu) = h(q) = -\sum_{i=0}^{d-1} q_i \log q_i.$$

It is a classical theorem of Sinai [5] that whenever $h(\mu) \geq h(\nu)$, there is a factor map from $B(p)$ to $B(q)$.

The second necessary condition for the existence of such a monotone factor is that μ **stochastically dominates** ν, which we now define. A **coupling** γ of two measures (X_1, σ_1) and (X_2, σ_2) is a measure on $X_1 \times X_2$ such that the marginal measures are σ_1 and σ_2:

$$\sigma_1(\cdot) = \gamma(\cdot \times X_2), \quad \sigma_2(\cdot) = \gamma(X_1 \times \cdot).$$

Coupling is a term from probability. In ergodic theory, the more usual notion is that of a joining. A **joining** of two dynamical systems (X_1, σ_1, T_1) and (X_2, σ_2, T_2) is a coupling of σ_1 and σ_2 which is invariant under the product map $T_1 \times T_2$ acting on $X_1 \times X_2$.

If there is a partial ordering $\geq$ defined on $X_1 \cup X_2$, then a coupling γ of σ_1 and σ_2 is **monotone** if

$$\gamma(\{(x_1, x_2) \colon x_1 \geq x_2\}) = 1.$$

We say that μ stochastically dominates ν ($\mu \succcurlyeq \nu$) if there is a monotone coupling of μ and ν. If μ and ν are preserved by maps T_1 and T_2 respectively, the existence of a monotone coupling of the measures implies the existence of a monotone joining of T_1 and T_2; just average the monotone coupling over orbits of $T_1 \times T_2$.

Example 1.1: For p and q probability measures on $A = \{0, \dots, d-1\}$, $p \succcurlyeq q$ exactly when

$$\sum_{i=j}^{d-1} p_i \geq \sum_{i=j}^{d-1} q_i \quad \text{for all } i = 0, \dots, d-1. \tag{1}$$

It is clear that this is a necessary condition for the existence of a monotone coupling of p and q. To show that (1) is sufficient to imply $p \succcurlyeq q$, we give a coupling of the two measures. For $i = 0, \dots, d-1$, let $r_i = \sum_{j=i}^{d-1} p_j$ and $s_i = \sum_{j=i}^{d-1} q_j$. Define a measure γ on $A \times A$ by

$$\gamma(i,j) = \ell((r_{i+1}, r_i) \cap (s_{j+1}, s_j))$$

where ℓ is Lebesgue measure. By (1), γ is a monotone coupling of p and q. This construction works for any discrete space with a total ordering. Strassen's 1965 paper [6] gives a related characterization of stochastic domination in more generality.

Any monotone factor map ϕ between two dynamical systems gives a monotone coupling (in fact, a monotone joining): $\gamma(C_1 \times C_2) = \sigma_1(C_1 \cap \phi^{-1}(C_2))$. Therefore, it is necessary to the existence of a monotone factor from $B(p)$ to $B(q)$ that $\mu \succcurlyeq \nu$. It is easy to see that $\mu \succcurlyeq \nu$ if and only if $p \succcurlyeq q$.

The factors we construct here will have the stronger property of being finitary codes. A **finitary code** is a measure-preserving factor between symbolic systems on a sequence space $A^{\mathbb{Z}}$ which is, up to sets of measure zero, continuous in the natural product topology on this space. Equivalently, a factor map ϕ: $A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ is finitary if for almost every $x \in A^{\mathbb{Z}}$, there is an integer $N = N(x)$ such that for any x' with $(x_{-N}, \dots, x_N) = (x'_{-N}, \dots, x'_N)$, $(\phi(x))_0 = (\phi(x'))_0$.

The main theorem in this paper is the following.

THEOREM 1.2: *If q is supported on two symbols and*

$$h(p) > h(q),$$

$$p \succcurlyeq q,$$

then there exists a finitary code ϕ: $B(p) \to B(q)$ which is monotone.

Note that we have weakened the necessary condition $h(p) \geq h(q)$ to $h(p) > h(q)$. Our theorem also only applies to cases where $B(q)$ has only two symbols.

This is a technical requirement which is there because when there are more symbols, the structure of the partial order $\leq$ on $A^{\mathbb{Z}}$ becomes too complicated for our argument, though there seems to be no inherent reason that Theorem 1.2 should not hold for systems with larger alphabets.

To illustrate the point that the two-symbol assumption does not seem to be necessary, we include in Section 8 the proof of a theorem of which the following is a special case.

THEOREM 1.3: *If q is any probability vector which gives probability $1/n$ to each of n symbols, then for any p with*

$$h(p) > h(q) = \log n,$$
$$p \succcurlyeq q,$$

there is a finitary code ϕ: $B(p) \to B(q)$.

In [2], Keane and Smorodinsky prove that there exists a finitary code from $B(p)$ to $B(q)$ whenever the entropy constraint is satisfied. Our proofs will follow their outline, but new ideas are needed to get the monotonicity result. They have also shown [3] that there is an invertible finitary code from $B(p)$ to $B(q)$ whenever $h(p) = h(q)$.

The question of whether there exists a monotone factor from $B(p)$ to $B(q)$ when $h(p) > h(q)$ and $p \succcurlyeq q$ was posed by Russ Lyons. Yuval Peres and Alexander Holroyd have asked the related question of whether there exists a monotone factor from a Poisson process with rate λ to another Poisson process with rate λ' whenever $\lambda > \lambda'$. In this case there is no entropy constraint because Poisson processes have infinite entropy. This question is answered positively in [1]. These questions are interesting in that they combine the interests of probabilists in monotone couplings (which have been used in studying percolation, for example) and the ideas of factors and finitary factors from ergodic theory.

2. Markers and skeletons

In this section, we set up some notation and outline the construction of our factor map ϕ. We begin by introducing the notion of markers in $B(p)$. Markers define for us boundaries between finite segments in our sequence space X, which is useful since we want ϕ to be a finitary code. For the rest of the paper (except the last section), p and q are probability vectors on A with $p \succcurlyeq q$, $h(p) > h(q)$, and q supported on exactly two symbols.

Remark 2.1: We will assume that $p_0, p_1, q_0, q_1 > 0$. These assumptions reduce the notation but are not necessary. The same argument will work if the roles of 0 and 1 in $B(q)$ are played by $i_0 < i_1$, the two nonzero entries of q, and the roles of 0 and 1 in $B(p)$ are played by any j_0 and j_1 in A with $p_{j_0}, p_{j_1} > 0$.

Fix an integer $k > 0$. The precise choice of k will be done later. A **marker** for $B(p)$ is a block

$$M = 0^{2k-1}1 = \underbrace{0, 0, \ldots 0}_{2k-1}, 1.$$

Since ϕ will be monotone, ϕ must take any marker to either 0^{2k} or $0^{2k-1}1$. In the more general case mentioned in Remark 2.1, there may be many possible monotone images for a marker sequence. In [2], the finitary code was constructed so that the word M always mapped to M. The restrictions imposed by monotonicity make it more convenient for us to allow all possible monotone images of markers.

Define a skeleton for $B(p)$ to be a sequence of the form

$$M^{n_0} \text{——} M^{n_1} \text{——} \cdots \text{——} M^{n_{t-1}} \text{——} M^{n_t} \quad \text{where } n_r < \min(n_0, n_t) \text{ for } 1 \le r \le t-1. \tag{2}$$

Here, —— stands for a stretch of the process with no markers. A skeleton s' is said to be a **subskeleton** of a skeleton s if

$$s = M^{n_0} \underbrace{\text{——}}_{l_1} M^{n_1} \underbrace{\text{——}}_{l_2} \cdots \underbrace{\text{——}}_{l_{t-1}} M^{n_{t-1}} \underbrace{\text{——}}_{l_t} M^{n_t} \tag{3}$$

and there exist $0 \le r < r' \le t$ such that

$$s' = M^{n_r} \underbrace{\text{——}}_{l_{r+1}} M^{n_{r+1}} \text{——} \cdots \underbrace{\text{——}}_{l_{r'}} M^{n_{r'}}.$$

Two subskeletons are said to be **disjoint** if they overlap at most in markers. A subskeleton s' is **maximal** in s if there does not exist a subskeleton s'' with $s' \subsetneq s'' \subsetneq s$. By the definition of a skeleton, either s has no proper subskeletons, or else it is covered by its maximal subskeletons, which are all disjoint.

Following Keane and Smorodinsky, we also inductively define an order on skeletons. A skeleton has **order** one if it has no proper subskeletons. A skeleton is of **order** n if it is not of order less than n but all of its proper subskeletons are of order less than n. It is not difficult to see that this notion is well-defined.

Let A_0^l denote the set of words of length l with symbols from A which do not contain a marker M. The space of all possible words filling in a skeleton s with structure

$$M^{n_0} \underbrace{\text{———}}_{l_1} M^{n_1} \underbrace{\text{———}}_{l_2} \cdots \underbrace{\text{———}}_{l_{t-1}} M^{n_{t-1}} \underbrace{\text{———}}_{l_t} M^{n_t}$$

will be

$$\mathcal{F}(s) := A_0^{l_1} \times M^{n_1} \times \cdots \times A_0^{l_t} \times M^{n_t}.$$

Such words are called **fillers** for s. Note that the markers are included in $\mathcal{F}(s)$ for $1 \le i \le t$, but the initial markers are not included; this is because these will be in the end markers for the previous skeleton. The space of fillers for s in $B(q)$ is then the set of all words dominated under the partial order by some $g \in \mathcal{F}(s)$:

$$\bar{\mathcal{F}}(s) := A^{l_1} \times (0^{2k-1} \times \{0,1\})^{n_1} \times \cdots \times A^{l_t} \times (0^{2k-1} \times \{0,1\})^{n_t}.$$

Elements of $\bar{\mathcal{F}}(s)$ are also called fillers.

We now briefly outline how the code will be constructed. To each point in $x \in B(p)$, we can associate an increasing sequence of recursively-defined skeletons as follows:

- Let $s_1(x)$ be the smallest skeleton in x containing the zeroth coordinate in its interior (i.e. not in one the marker sequences at either end of the skeleton).
- Once $s_t(x)$ has been defined, let $s_{t+1}(x)$ be the next-smallest skeleton in x containing $s_t(x)$ in its interior.

It is clear that this whole sequence of skeletons exists for a point with probability one. Also note that when this sequence exists, each coordinate is covered by all $s_t(x)$ with t sufficiently large.

For each skeleton s, $\mathcal{F}(s)$ is the set of all possible words in $B(p)$ that can fill in the skeleton. $\bar{\mathcal{F}}(s)$ is the set of possible images for elements of $\mathcal{F}(s)$ under a monotone map. As in [2], we will use a marriage lemma which allows us to define partial maps from a subset of $\mathcal{F}(s)$ to $\bar{\mathcal{F}}(s)$. Once the image of some $g \in \mathcal{F}(s)$ is fixed, the extension of g in a higher-order skeleton containing s as a subskeleton will be assigned the same image on s, so that values cannot be changed at a later stage of the construction.

We prove that these choices may be made in a monotone way. Our method for defining the partial maps of fillers is similar to that used in [2]. However, our work includes new ideas to deal with the fact that we have to be careful with the points where the partial map is not defined in order to ensure that a monotone image for these points will be available at a later stage of the construction.

We define ϕ as follows. The point x has a certain filler for $s_1(x)$. If that filler has an image defined by the marriage lemma, then that image gives the value of $\phi(x)$ on s_1, including the value at zero. Otherwise, one looks at $s_2(x)$, $s_3(x)$, and so on until the image of the filler is defined. The construction is such that the image of the filler for $s_t(x)$ will exist for t sufficiently large, almost surely. For all t for which the image exists for x, the image will give the same value for $\phi(x)_0$. Such a procedure must result in a code which is finitary and which commutes with the shift map. It is also simple given the construction to show that the image of ϕ is indeed $B(q)$.

3. Societies

The partial maps we define from $\mathcal{F}(s)$ to $\bar{\mathcal{F}}(s)$ will be defined using a marriage lemma, Lemma 6.2, which is adapted from the one used in [2]. In this section, we give the relevant definitions to formulate and prove this lemma. We start by defining societies.

Let (U,ρ), (V,σ) be finite measure spaces with U a finite set. In keeping with the usual terminology of a marriage lemma, the elements of U will sometimes be referred to as boys and the elements of V are girls. In this paper, boys will be elements of $\bar{\mathcal{F}}(s)$ (or a subset of $\bar{\mathcal{F}}(s)$) for some skeleton s and girls will be elements of $\mathcal{F}(s)$ (or a subset thereof). The measures we will use on these spaces will be given in Section 4.

A **society** from (U,ρ) to (V,σ) is a map $S\colon U \to 2^V$ such that for any $B \subseteq U$,

$$\rho(B) \le \sigma(S(B)) \quad \text{where } S(B) = \bigcup_{b\in B} S(b).$$

A society R is a **refinement** of S ($R \subseteq S$) if $R(b) \subseteq S(b)$ for all $b \in U$. A society S is **monotone** if there is a partial order $\le$ defined on $U \cup V$ and $b \le g$ for all $g \in S(b)$.

There is a relationship between societies and couplings which we will exploit. This relationship is given by the following remark and lemma, which can be found in Section 6.5 of Petersen's book, [4].

Remark 3.1: Any coupling of two probability measures on finite sets U and V gives rise to a society, as follows. Let γ be a coupling of (U,ρ) and (V,σ). Then we can define a map $S_\gamma : U \to 2^V$ by

$$S_\gamma(b) = \{g \in V : \gamma(b,g) > 0)\}.$$

Then for any $B \subseteq U$,

$$\begin{aligned}\rho(B) &= \gamma(B \times V) \\ &= \sum_{b \in B} \sum_{g \in S_\gamma(b)} \gamma(b,g) \leq \sum_{g \in S_\gamma(B)} \sum_{b \in U} \gamma(b,g) \\ &= \gamma(U \times S_\gamma(B)) = \sigma(S_\gamma(B)).\end{aligned}$$

Therefore, S_γ is a society. The following lemma gives a partial converse for this fact. See Petersen [4] for the proof.

LEMMA 3.2: *For any society S from (U, ρ) to (V, σ), there is a coupling γ of ρ and σ such that S_γ is a refinement of S.*

Because of these results, we can alternate between societies and couplings depending on which is more useful, so long as we are willing to pass to refinements of societies.

The following lemma shows that the product of two societies is also a society. Of course, this implies that the product of n societies is also a society. The proof is trivial.

LEMMA 3.3: *Let S_i be a society from (U_i, ρ_i) to (V_i, σ_i) for $i = 1, 2$. Then the map $S(b_1, b_2) = S_1(b_1) \times S_2(b_2)$ is a society from $(U_1 \times U_2, \rho_1 \times \rho_2)$ to $(V_1 \times V_2, \sigma_1 \times \sigma_2)$.*

4. A special coupling

Let s be a skeleton. Our object in this section is to construct a society from $\bar{\mathcal{F}}(s)$ to $\mathcal{F}(s)$ (we leave the measures unspecified for the moment) which is

1. monotone and
2. controls how often certain "good" and "bad" boys in $\bar{\mathcal{F}}(s)$ are allowed to relate to the same girls in $\mathcal{F}(s)$.

These societies will be defined by a special coupling γ of μ and ν, which we now define.

To make γ, we will couple the coordinates independently one at a time near the ends of markers in order to ensure that the coupling is independent on disjoint skeletons. Between markers, we want to try to reduce the number of girls each boy can couple to. To accomplish these two goals, γ is defined using two couplings, $\bar{\pi}$ and $\bar{\gamma}$, of μ and ν restricted to finite intervals.

Let $\bar{\pi}$ be a monotone coupling of p and q. Then a natural monotone coupling of $B(p)$ and $B(q)$ can be formed by taking $\pi = \prod_{t \in \mathbb{Z}} \bar{\pi}$. However, this coupling does not adequately control how many boys a girl will know.

The second coupling we define, $\bar{\gamma}$, will give us this control. $\bar{\gamma}$ is a monotone coupling of $B(p)$ and $B(q)$ restricted to a window of length k, where $2k$ is the length of a marker. First, partition A^k into sets

$$L_a = \{b \in A^k : b \text{ has exactly } a \text{ symbols } \geq 1\} \quad \text{for } 0 \leq a \leq k.$$

Let $\bar{\mu}$ and $\bar{\nu}$ be the measures induced by μ and ν on $\{L_a\}_{0 \leq a \leq k}$ respectively. We give an explicit coupling γ^* of $\bar{\mu}$ and $\bar{\nu}$. Consider

$$1 = r_0 > r_1 > \cdots > r_{k+1} = 0 \quad \text{where } r_a = \bar{\mu}(\{L_{a'}\colon a' \geq a\}),$$
$$1 = s_0 > s_1 > \cdots > s_{k+1} = 0 \quad \text{where } s_a = \bar{\nu}(\{L_{a'}\colon a' \geq a\}).$$

Now let ℓ be Lebesgue measure and define

$$\gamma^*(L_a, L_{a'}) = \ell((r_{a+1}, r_a) \cap (s_{a'+1}, s_{a'})).$$

Clearly, γ^* is a coupling of $\bar{\mu}$ and $\bar{\nu}$. Since $p \succcurlyeq q$, if $\gamma^*(L_a, L_{a'}) > 0$, then $a \geq a'$.

The coupling $\bar{\gamma}$ is now constructed by selecting the levels L_a and $L_{a'}$ from γ^*, then making all other choices as independently as possible while assuring monotonicity. Explicitly, for any pair $(g, b) \in A^k \times A^k$,

$$\bar{\gamma}(g, b) = \gamma^*(L(g), L(b)) \cdot \mu(g \mid L(g)) \cdot \nu(b \mid L(b) \cap \{b'\colon b' \leq g\}).$$

where $L(g)$ is the set L_a such that $g \in L_a$ and similarly for $L(b)$.

LEMMA 4.1: *$\bar{\gamma}$ is a monotone coupling of μ and ν on $A^k \times A^k$.*

Proof: It is clear that $\bar{\gamma}$ is monotone. We just have to show that it has the right marginal measures. The fact that the first marginal of $\bar{\gamma}$ is μ is easy and we leave it to the reader. We now prove that the second marginal measure of $\bar{\gamma}$ is ν. Let $b \in \{0,1\}^k$ with $L(b) = L_j$. Then

$$\bar{\gamma}(A^k \times b) = \sum_{g \in A^k} \gamma^*(L(g), L(b)) \cdot \mu(g \mid L(g)) \cdot \nu(b \mid L(b) \cap \{b'\colon b' \leq g\})$$

$$(4) \qquad = \sum_{i=j}^{k} \gamma^*(L_i, L_j) \sum_{\substack{g\colon g \geq b, \\ L(g) = L_i}} \mu(g \mid L_i) \nu(b \mid L_j \cap \{b'\colon b' \leq g\})$$

$$(5) \qquad = \sum_{i=j}^{k} \gamma^*(L_i, L_j) \binom{i}{j}^{-1} \mu(\{g\colon g \geq b\} \mid L_i)$$

$$(6) \qquad = \sum_{i=j}^{k} \gamma^*(L_i, L_j) \binom{i}{j}^{-1} \binom{k-j}{i-j} \binom{k}{i}^{-1}$$

$$(7) \qquad = \sum_{i=j}^{k} \gamma^*(L_i, L_j) \nu(b \mid L_j) = \nu(L_j) \nu(b \mid L_j) = \nu(b).$$

Note that to go from (4) to (5) and from (6) to (7) above, we used the fact that all words in the set L_j have the same measure under ν, which follows from the two-symbol assumption for q. ∎

The main property of $\bar{\gamma}$ that we will use is that if a, a' are sufficiently far apart, then $\{g: \bar{\gamma}(g \times L_a) > 0\} \cap \{g: \bar{\gamma}(g \times L_{a'}) > 0\} = \emptyset$. In other words, if $S_{\bar{\gamma}}$ is the society from (A^k, ν) to (A^k, μ) given by $\bar{\gamma}$, then any two boys with $b \in L_a$ and $b' \in L_{a'}$ will know no girls in common under $S_{\bar{\gamma}}$. This property is formally stated in the next lemma.

LEMMA 4.2: *For any $1 < n^* \leq k$, if we set*

$$a^* = a^*(n^*) = \max\{a: \gamma^*(L_{n^*}, L_a) > 0\},$$

then for any $a < a^* < a'$,

$$S_{\bar{\gamma}}(L_a) \cap S_{\bar{\gamma}}(L_{a'}) = \emptyset.$$

Proof: This follows immediately from the construction. If $a < a^* < a'$, then $S(L_a) \subseteq \bigcup_{n \leq n^*} L_n$ and $S(L_{a'}) \subseteq \bigcup_{n > n^*} L_n$. ∎

We use $\bar{\gamma}$ and $\bar{\pi}$ to define a joining γ between $B(p)$ and $B(q)$. Generate a sequence of pairs of random variables $Z_1, Z_2, \ldots$, with each Z_i taking values in $A \times A$ as follows. Let the vector $(Z_1, \ldots, Z_k)$ be chosen from the distribution $\bar{\gamma}$. Independently choose $(Z_{nk+1}, \ldots, Z_{(n+1)k})$ from $\bar{\gamma}$ for $n = 1, 2, \ldots$ until an n is reached for which $Z_{(n-1)k+1} = \cdots = Z_{nk} = (0,0)$ (so that n is the first time a sampling from $\bar{\gamma}$ results in all zeros). At this point, start choosing $Z_{nk+1}, Z_{nk+2}, \ldots$ independently from $\bar{\pi}$. Do this until the first time the first coordinate of Z_i is at least one. Return to generating $(Z_{i+1}, \ldots, Z_{i+k})$, from $\bar{\gamma}$ and so on. Continue to oscillate between generating the Z_i from $\bar{\gamma}$ and $\bar{\pi}$ as $i \to \infty$.

Let $\dot{\gamma}$ be the distribution of $Z_1, Z_2, \ldots$, which is clearly a coupling of $B(p)$ and $B(q)$ on $\mathbb{Z}^+$. However, $\dot{\gamma}$ is not stationary, so let

$$\ddot{\gamma} = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} T^{-i} \dot{\gamma},$$

which will be stationary and in fact is an ergodic joining of $B(p)$ and $B(q)$ on $\mathbb{Z}^+$. Finally, let γ be the natural extension of $\ddot{\gamma}$ to an ergodic joining of $B(p)$ and $B(q)$ on all of $\mathbb{Z}$.

From the construction above, it is obvious that γ is a monotone joining of μ and ν. We now show that γ conditioned on its first coordinate has a convenient product structure. For an interval $K = [i, j)$ and a point $x \in A^{\mathbb{Z}}$, let $x_K = (x_i, \dots, x_{j-1}) \in A^k$. For $x \in X$, define a sequence $\mathbf{K}^x = (\dots, K_{-1}^x, K_0^x, K_1^x, \dots)$ of disjoint intervals of integers covering $\mathbb{R}$:

1. If $x_{[j-2k+1,j]} = M$, then the interval $[j+1, j+k] \in \mathbf{K}^x$.
2. If $K_i \in \mathbf{K}^x$, then we can find the length of K_{i+1}:
   - for $|K_i| = k$: $|K_{i+1}| = k$ if $x_{K_i} \neq 0^k$; otherwise, $|K_{i+1}| = 1$,
   - for $|K_i| = 1$: $|K_{i+1}| = k$ if $x_{K_i} = 1$; otherwise, $|K_{i+1}| = 1$.

Since a marker has the form $0^{2k-1}1$, any interval $K \in \mathbf{K}^x$ containing the endpoint of a marker must have $|K| = 1$ and $x_K = 1$, so a new interval of length k must be started after the end of each marker and thus the sequence above is well-defined. $\mathbf{K}^x$ is then defined for μ-almost every x and the intervals are chosen in such a way that if $\{B_t \subset A^{|K_t^x|}\}_{t\in\mathbb{Z}}$ and $\mathbf{B} = \{y \in Y : y_{K_t^x} \in B_t\}$, then

$$\gamma(x \times \mathbf{B} \mid x \times Y) = \prod_{t\in\mathbb{Z}} (\mathbf{1}_{|K_t^x|=k} \bar{\gamma}(x_{K_t^x} \times B_t \mid x_{K_t^x} \times A^k) + \mathbf{1}_{|K_t^x|=1} \bar{\pi}(x_{K_t^x} \times B_t \mid x_{K_t^x} \times A)) \text{ a.s.} \tag{8}$$

What this says is that on a k-segment $K \in \mathbf{K}^x$, we always sample from $\bar{\gamma}$ conditioned on x_K and on a 1-segment $K' \in \mathbf{K}^x$, we sample from $\bar{\pi}$, given $x_{K'}$. All of these samplings are independent given x. This gives the product structure above.

Similarly, an element $g \in \mathcal{F}(s)$ gives a unique finite sequence of intervals $\mathbf{K}^g = \mathbf{K}^x|_s$ for any $x|_s = g$. After the end of a marker, the coupling always samples first from $\bar{\gamma}$, conditioned on x, so K^g always starts with a k-segment.

Now, let $\mathbf{K}$ be a finite sequence of intervals of length 1 or k, starting with an interval of length k, which covers and is consistent with a skeleton s (excluding its initial marker sequence). Let $\bar{l}$ be the length of s, excluding its initial markers. Define

$$\mathcal{F}(s, \mathbf{K}) := \{g \in A^{\bar{l}} : g \in \mathcal{F}(s) \text{ and } \mathbf{K}^g = \mathbf{K}\},$$
$$\gamma_{s,\mathbf{K}}(\cdot) := \gamma(\cdot \mid \mathcal{F}(s, \mathbf{K}) \times \bar{\mathcal{F}}(s)),$$
$$\bar{\mathcal{F}}(s, \mathbf{K}) := \{b \in A^{\bar{l}} : \exists g \in \mathcal{F}(s, \mathbf{K}) \text{ with } \gamma_{s,\mathbf{K}}(g, b) > 0\}.$$

Then $\gamma_{s,\mathbf{K}}$ is a measure on $\mathcal{F}(s, \mathbf{K}) \times \bar{\mathcal{F}}(s, \mathbf{K})$. Set

$$\mu_{s,\mathbf{K}}(\cdot) := \mu(\cdot \mid \mathcal{F}(s, \mathbf{K})) = \gamma_{s,\mathbf{K}}(\cdot \times \bar{\mathcal{F}}(s, \mathbf{K})),$$
$$\nu_{s,\mathbf{K}}(\cdot) := \gamma_{s,\mathbf{K}}(\mathcal{F}(s, \mathbf{K}) \times \cdot).$$

Under this notation, $\gamma_{s,\mathbf{K}}$ is a coupling of $\mu_{s,\mathbf{K}}$ and $\nu_{s,\mathbf{K}}$. It is an important fact that if $s_1, \dots, s_t$ are the maximal subskeletons of s and $\mathbf{K}_i = \mathbf{K}|_{s_i}$ for

$i = 1, \dots, t$, then by (8):

$$\gamma_{s,\mathbf{K}} = \gamma_{s_1,\mathbf{K}_1} \times \cdots \times \gamma_{s_t,\mathbf{K}_t}. \tag{9}$$

Let $Q_{s,\mathbf{K}} = S_{\gamma_{s,\mathbf{K}}}$ be the society given by $\gamma_{s,\mathbf{K}}$ from $(\bar{\mathcal{F}}(s,\mathbf{K}), \nu_{s,\mathbf{K}})$ to $(\mathcal{F}(s,\mathbf{K}), \mu_{s,\mathbf{K}})$, which is by definition monotone. We will apply our Marriage Lemma to these societies to define ϕ.

It is also important that $\mu_{s,\mathbf{K}}$ and $\nu_{s,\mathbf{K}}$ are disintegrations of μ and ν. More explicitly, for $x \in X$, let $s(x) = \{s_t^{(1)}(x)\}_{t\in\mathbb{Z}}$ be the sequence of order-one skeletons covering $\mathbb{Z}$ defined by x. This sequence is defined almost surely. For any interval I, write $\mathbf{K}^x(I) = \mathbf{K}^x|_I$. Set

$$\mu_{s(x),\mathbf{K}^x} := \prod_{t\in\mathbb{Z}} \mu_{s_t^{(1)}(x),\mathbf{K}^x(s_t^{(1)}(x))} \quad \text{and} \quad \nu_{s(x),\mathbf{K}^x} := \prod_{t\in\mathbb{Z}} \nu_{s_t^{(1)}(x),\mathbf{K}^x(s_t^{(1)}(x))}.$$

Since μ is a product measure, we have $\mu = \int \mu_{s(x),\mathbf{K}^x} d\mu(x)$. The fact that γ is a coupling of μ and ν and (9) then imply that $\nu = \int \nu_{s(x),\mathbf{K}^x} d\mu(x)$. Partition X into sets of the form

$$X(x) = \{x' \in X : s(x') = s(x) \text{ and } \mathbf{K}^{x'} = \mathbf{K}^x\}.$$

If we can define a monotone map $\phi_{X(x)}$ from $X(x)$ to Y which takes $\mu_{s(x),\mathbf{K}^x}$ to $\nu_{s(x),\mathbf{K}^x}$ almost surely, then $\phi(x) = \phi_{X(x)}(x)$ will be the desired monotone factor. This is our aim.

5. Good and bad fillers

As in [2], we will distinguish between good (typical) fillers and bad fillers for a skeleton s. However, because we are constructing a monotone code, we cannot treat all bad boys alike. While making the partial assignments, we need to make sure to leave enough space for all of the bad boys to get later assignments which are monotone. This means that we cannot lump bad boys together indiscriminantly, because the sets of girls they can associate to are not the same. For this reason, our definition of good and bad boys are more complicated than in [2].

Fix $\epsilon > 0$, to be chosen later. Our choice of k sufficiently large, to be used in the definitions of markers and $\bar{\gamma}$, will then be made depending on ϵ. The parameters ϵ and k will be used to define our sets of bad girls and boys.

Let s be a skeleton and $\mathbf{K}$ a set of intervals consistent with s. Let $\bar{l} = \sum_{i=1}^t (l_i + 2kn_i)$, the total length of s except for its initial markers. Partition

K into four sets

$$\mathbf{K}_1 = \{K_i \in \mathbf{K}\colon |K_i| = 1, |K_{i+1}| = 1\},$$
$$\mathbf{K}_2 = \{K_i \in \mathbf{K}\colon |K_i| = 1, K_i \notin \mathbf{K}_1\},$$
$$\mathbf{K}_3 = \{K_i \in \mathbf{K}\colon |K_i| = k, |K_{i+1}| = 1\},$$
$$\mathbf{K}_4 = \{K_i \in \mathbf{K}\colon |K_i| = k, |K_{i+1}| = k\}.$$

By the construction of $\mathbf{K}^g$, $\mathbf{K} = \mathbf{K}^g$ determines g_K for all $K \in \mathbf{K} \setminus \mathbf{K}_4$:

$$\begin{aligned}\mathcal{F}(s,\mathbf{K}) = \{g \in A^{\bar{l}} :& g_K = 0 \text{ for } K \in \mathbf{K}_1, \quad g_K = 1 \text{ for } K \in \mathbf{K}_2,\\ & g_K = 0^k \text{ for } K \in \mathbf{K}_3, \quad g_K \in A^k \setminus \{0^k\} \text{ for } K \in \mathbf{K}_4\}.\end{aligned} \tag{10}$$

Set $l = k \cdot (\#\mathbf{K}_4)$, the number of coordinates of s covered by intervals $K \in \mathbf{K}$ for which there is a choice for x_K. Clearly, $l < \bar{l}$ for all $(s, \mathbf{K})$.

We say an element $g \in \mathcal{F}(s, \mathbf{K})$ is a **good filler** or a **good girl** if

$$\mu_{s,\mathbf{K}}(g) \leq 2^{-(h(p)-\epsilon)l}.$$

The definition of a good boy in $\bar{\mathcal{F}}(s, \mathbf{K})$ is more complicated, though it will be related to the set of $b \in \bar{\mathcal{F}}(s, \mathbf{K})$ with $\nu_{s,\mathbf{K}}(b) \geq 2^{-(h(q)+\epsilon)\bar{l}}$. First we concentrate on elements of $\{0,1\}^k$. An atom in $\{0,1\}^k$ is called a **k-segment**. Let

$$H = \{b \in \{0,1\}^k : \nu(b) \leq 2^{-(h(q)+\epsilon)k}\}.$$

The points in H will form the core of the set of bad k-segments, which will in turn be used to define bad boys $b \in \bar{\mathcal{F}}(s, \mathbf{K})$.

Let $S_{\bar{\gamma}}$ be the society from $(\{0,1\}^k, \nu)$ to (A^k, μ) induced by the coupling $\bar{\gamma}$. We want to make use of the nice properties of $S_{\bar{\gamma}}$, and in particular Lemma 4.2, to control how much boys in H and $\{0,1\}^k \setminus H$ can associate to the same girls in A^k.

Note that because of the two-symbol assumption on q, for some a',

$$H = \begin{cases} \bigcup_{a \geq a'} L_a & \text{if } q_0 > 1/2, \\ \emptyset & \text{if } q_0 = 1/2, \\ \bigcup_{a \leq a'} L_a & \text{if } q_0 < 1/2. \end{cases}$$

Let

$$n^* = \begin{cases} \min\{n\colon \bar{\gamma}(L_n \times H) > 0\} & \text{if } q_0 > 1/2, \\ \max\{n\colon \bar{\gamma}(L_n \times H) > 0\} & \text{if } q_0 < 1/2, \end{cases}$$
$$a^* = \begin{cases} \max\{a\colon \bar{\gamma}(L_{n^*-1}, L_a) > 0\} & \text{if } q_0 > 1/2, \\ \min\{a\colon \bar{\gamma}(L_{n^*+1}, L_a) > 0\} & \text{if } q_0 < 1/2. \end{cases}$$

We say that a k-segment $b \in \{0,1\}^k$ is **bad** if it lies in the set

$$B = \begin{cases} \bigcup_{a \geq a^*} L_m & \text{if } q_0 > 1/2, \\ \emptyset & \text{if } q_0 = 1/2, \\ \bigcup_{a \leq a^*} L_m & \text{if } q_0 < 1/2. \end{cases}$$

Clearly, $H \subset B \setminus L_{a^*}$.

Now if $b \in \bar{\mathcal{F}}(s, \mathbf{K})$, we say that b is a good boy if most of its k-segments are not bad. Explicitly, b is a **good boy** if

$$\frac{\#\{K \in \mathbf{K}_4 \colon b_K \in B\}}{\#\{K \in \mathbf{K}_4\}} < \epsilon.$$

Define $B^* \subseteq \bar{\mathcal{F}}(s, \mathbf{K})$ by

$$B^* = \left\{b \colon b \text{ is not a good boy but } \frac{\#\{K \in \mathbf{K}_4 : b_K \in B \setminus L_{a^*}\}}{\#\{K \in \mathbf{K}_4\}} < \epsilon\right\}.$$

This set is important because, as we will see in the proof of Lemma 6.2, if $b \in \bar{\mathcal{F}}(s, \mathbf{K})$ is bad and there exist a good boy $b' \in \bar{\mathcal{F}}(s, \mathbf{K})$ and a girl $g \in \mathcal{F}(s, \mathbf{K})$ such that $g \in Q_{s,\mathbf{K}}(b) \cap Q_{s,\mathbf{K}}(b')$, then $b \in B^*$. In other words, the boys in B^* are the only bad boys who know any girls in common with good boys under $Q_{s,\mathbf{K}}$. Along these lines, define

$$U' = \{b \in \bar{\mathcal{F}}(s, \mathbf{K}) \colon b \text{ is good}\} \cup B^*.$$

In the rest of this section, we prove technical lemmas showing that the definitions above are reasonable; that most girls are good, that most k-segments are good, that for most large skeletons, most boys are good boys, and that the number of boys in U' is not too large.

Recall that $s_1(x) \subsetneq s_2(x) \subsetneq \cdots$ is a sequence of skeletons in x covering the origin and with $\cup_{t \in \mathbb{N}} s_t(x)$ covering $\mathbb{Z}$, as defined in Section 2.

Lemma 5.1: *For all $\epsilon > 0$ and all k sufficiently large, for a.e. $x \in X$, $x|_{s_t}$ is a good girl in $\mathcal{F}(s_t(x), \mathbf{K}^x|_{s_t(x)})$ when t is sufficiently large.*

Proof: Fix x and for each $t \in \mathbb{Z}$, set $s_t = s_t(x)$, $\mathbf{K}(t) = \mathbf{K}^x|_{s_t}$, and $x_t = x|_{s_t}$. Then because the first marginal of $\tilde{\gamma}$ is μ and by (10),

$$\mu_{s_t, \mathbf{K}(t)}(x_t) = \prod_{K \in \mathbf{K}_4(t)} \mu(x_K \mid A^k \setminus \{0^k\}).$$

Choose k large enough so that

$$\mu(\{g \in A^k \colon \mu(g) \leq 2^{-(h(p) - \epsilon/2)k}\}) > 1 - \frac{\epsilon}{4 \max\{1, h(p)\}}, \tag{11}$$

which exists by the Shannon–McMillan–Breiman theorem. Then for k large,

$$\mu(\{g \in A^k \colon \mu(g) \le 2^{-(h(p)-\epsilon/2)k}\} \mid A^k \setminus \{0^k\}) \\ > 1 - \frac{\epsilon}{4\max\{1, h(p)\}} - p_0^k > 1 - \frac{\epsilon}{2\max\{1, h(p)\}}.$$

Almost surely, as $t \to \infty$, $\#\mathbf{K}_4(t) \to \infty$. When this is the case, the law of large numbers implies that almost surely the proportion of $K \in \mathbf{K}_4(t)$ with x_K not in the set defined in (11) is less than $\epsilon/2\max\{1, h(p)\}$ for all t sufficiently large. For such values of t, we have

$$\mu_{s_t,\mathbf{K}(t)}(x_t) \le 2^{-(h(p)-\epsilon/2)k\cdot(\#\mathbf{K}_4)(1-\epsilon/(2\max\{1,h(p)\}))} \le 2^{-(h(p)-\epsilon)l}.$$

Thus for almost all x, if t is sufficiently large then x_t is a good girl in $\mathcal{F}(s_t(x), \mathbf{K}(t))$. ∎

Lemma 5.2: *For any $\epsilon > 0$, if k is sufficiently large then $\nu(B) < \epsilon$.*

Proof: If $q_0 = 1/2$, then $\nu(B) = 0$. We do the proof for the case where $q_0 > 1/2$. The case where $q_0 < 1/2$ is similar. When $q_0 > 1/2$,

$$B = H \cup L_{a^*} \cup L_{a^*+1} \cup \cdots \cup L_{a'-1}$$

where a' is the smallest integer with $L_{a'} \subseteq H$. By the Shannon–McMillan–Breiman Theorem, $\nu(H) < \epsilon/2$ for k sufficiently large.

By the definition of a^* and n^* and the construction of γ^*, either $a^* = a' - 1$ or

$$\nu(L_{a^*} \cup L_{a^*+1} \cup \cdots \cup L_{a'-1}) \\ = \gamma^*((L_{a^*} \cup \cdots \cup L_{a'-1}) \times (L_{n^*} \cup L_{n^*-1})) \le \mu(L_{n^*} \cup L_{n^*-1}).$$

We need to show then that L_{n^*} and L_{n^*-1} have small size under μ and that L_{a^*} has small measure under ν. Let $\tilde{p}_1 = \sum_{i=1}^{k} p_i$ be the probability under p of being at least one and let $\tilde{p}_0 = 1 - \tilde{p}_1$. The set L_a with largest size will have either $a = \lceil \tilde{p}_1 k \rceil$ or $a = \lfloor \tilde{p}_1 k \rfloor$ since the number of symbols which are at least 1 in a k segment has a binomial distribution with mean $\tilde{p}_1 k$. Then for any n,

$$\mu(L_n) = \binom{k}{n} \tilde{p}_0^{(k-n)} \tilde{p}_1^n \le \binom{k}{\tilde{p}_1 k} \tilde{p}_0^{\tilde{p}_0 k} \tilde{p}_1^{\tilde{p}_1 k} \le \frac{C}{\sqrt{k}},$$

where C is a constant depending on p and the last inequality follows from Stirling's formula. The same reasoning shows that $\nu(L_{a^*}) < C'/\sqrt{k}$. Thus, as $k \to$

∞ the maximum size of a level goes to zero and $\max\{\mu(L_{n^*} \cup L_{n^*-1}), \nu(L_{a^*})\} < \epsilon/2$ for k sufficiently large. Therefore, $\nu(B) < \epsilon$. ∎

Next, we prove that for large skeletons, most boys are good.

LEMMA 5.3: *For any ϵ, if k is chosen sufficiently large, then for almost every x,*

$$\lim_{t\to\infty} \nu_{s_t(x),\mathbf{K}^x|_{s_t(x)}}(\{b \in \bar{\mathcal{F}}(s_t(x), \mathbf{K}^x|_{s_t(x)}) \colon b \text{ is good}\}) = 1. \tag{12}$$

Proof: Fix x and for each $t \in \mathbb{Z}$, set $s_t = s_t(x)$ and $\mathbf{K}(t) = \mathbf{K}^x|_{s_t}$. Under $\nu_{s_t,\mathbf{K}(t)}$, the values of b on each $K \in \mathbf{K}_4(t)$ are i.i.d. with distribution

$$\nu_*(\cdot) = \bar{\gamma}(A^k \times \cdot \mid (A^k \times A^k) \setminus \{(0,0)^k\}).$$

We can estimate the size of the set B of bad k-segments:

$$\begin{aligned}\nu_*(B) &= \bar{\gamma}(A^k \times B| \,(A^k \times A^k) \setminus \{(0,0)^k\}) \le \frac{\bar{\gamma}(A^k \times B)}{1 - \bar{\gamma}(\{(0,0)^k\})}\\ &= \frac{\nu(B)}{1-p_0^k} \quad \text{since } \bar{\gamma}(\{(0,0)^k\}) = p_0^k \text{ by monotonicity}\\ &< \epsilon \quad \text{by Lemma 5.2 when } k \text{ is large.}\end{aligned}$$

A boy $b \in \bar{\mathcal{F}}(s_t, \mathbf{K}(t))$ is good when the fraction of $K \in \mathbf{K}_4(t)$ with $b_K \in B$ is less than ϵ. Since the values of b_K are i.i.d. and the probability that $b_K \in B$ is less than ϵ, the law of large numbers implies (12). ∎

The last results in this section establish a bound on the number of elements of U' for most pairs $(s, \mathbf{K})$. Since the definition of U' depends only on the values of boys on $\mathbf{K}_4$, we need to show that most of the time, s is mostly covered by $\mathbf{K}_4$. We say that a pair $(s, \mathbf{K})$ is **good** if $l \ge \bar{l}(1-\epsilon)$. In other words, if $(s, \mathbf{K})$ is good then $\mathbf{K}_4$ covers all but at most ϵ of the coordinates of s.

LEMMA 5.4: *For all $\epsilon > 0$ and k sufficiently large, we have that for almost all $x \in X$, the pairs $(s_t(x), \mathbf{K}^x|_{s_t(x)})$ are good for all t sufficiently large.*

Proof: By (10), for any $K \in \mathbf{K}^x$ with $K \notin \mathbf{K}_4^x$, K is contained in an interval I of length at least $k+1$ such that $x_I = 0^k1$. For k large, this is an event with probability less than ϵ. Then by the ergodic theorem, the fraction of coordinates i in s_t with $i \in K \in \mathbf{K}_4$ is almost surely at least $1-\epsilon$ for all t sufficiently large. ∎

We are now ready to estimate the size of U' for good $(s, \mathbf{K})$. Let $h_{ent}(\alpha) := -\alpha \log \alpha - (1-\alpha)\log(1-\alpha)$ be the entropy function.

LEMMA 5.5: *Let $(s, \mathbf{K})$ be good with total length $\bar{l}$. The number of boys in $U' \subseteq \bar{\mathcal{F}}(s, \mathbf{K})$ is at most $2^{\bar{l}(h(q)+3\epsilon+h_{ent}(\epsilon))}$.*

Proof: For any boy $b \in U'$, the fraction of $K \in \mathbf{K}_4$ with $b_K \in H \subseteq B \setminus L_{a^*}$ is less than $1-\epsilon$. Recall that the definition of H is that it consists of all $x \in A^k$ with $\nu(x) \leq 2^{-(h(q)+\epsilon)k}$. Therefore, there are at most $2^{(h(q)+\epsilon)k}$ elements in $A^k \setminus H$. Since $(s, \mathbf{K})$ is good, $\mathbf{K}_4$ covers all but ϵ of s. Therefore, we have

$$|U'| \leq \binom{\#\mathbf{K}_4}{\epsilon(\#\mathbf{K}_4)} 2^{2\epsilon\bar{l}} 2^{(h(q)+\epsilon)\bar{l}} \leq 2^{(h(q)+3\epsilon+h_{ent}(\epsilon))\bar{l}},$$

where the last inequality follows from Stirling's formula. ∎

6. Assignments for fillers

In this section, we give the societies which will be used to define the partial assignments for fillers at each stage of the construction of the finite code. These societies are defined using a slight variation of the Marriage Lemma from [2].

Let S be a society from (U, ρ) to (V, σ) and assume that U and V are both finite sets. Let $U' \subseteq U$. Define

$$\pi_{U'}(S) := \#\{g \in V\colon \exists b_1, b_2 \in U', b_1 \neq b_2, \text{ with } g \in S(b_1) \cap S(b_2)\}.$$

When $U' = U$, Keane and Smorodinsky call this number the promiscuity number of S, which is written $\pi(S)$, because it is the number of girls who know more than one boy under S.

LEMMA 6.1: *Let (U, ρ), (V, σ) be finite measure spaces with U and V both finite. Let S be a society from U to V and let $U' \subseteq U$. Then there is a refinement R of S such that $\pi_{U'}(R) \leq |U'| - 1$.*

Proof: Take R to be a society more refined than S with $\pi(R)$ minimal. This can be done since there are only finitely many societies from U to V.

A **cycle** is a sequence $b_1, \dots, b_t$ of distinct boys and a sequence $g_1, \dots, g_t$ of distinct girls such that for all $1 \leq j \leq t$,

$$g_j \in R(b_j) \cap R(b_{j+1}), \quad \text{with } b_{t+1} := b_1.$$

In Theorem 11 of [2], it is shown that R has no cycles. For completeness we include the proof, as given by Petersen in [4].

Suppose there is a cycle in R. By the minimality of R, there is a coupling λ of σ and ρ such that $S_\lambda = R$. Let $\alpha = \min_{1\le j\le t} \lambda(g_j, b_j)$ and define a new coupling:

$$\lambda^*(g,b) = \begin{cases} \lambda(g_j,b_j) - \alpha & \text{if } g = g_j, b = b_j, \\ \lambda(g_j,b_{j+1}) + \alpha & \text{if } g = g_j, b = b_{j+1}, \\ \lambda(g,b) & \text{otherwise.} \end{cases}$$

Then λ^* is another coupling of σ and ρ, and the induced society has a smaller promiscuity number than R. This gives a contradiction. Therefore, R has no cycles.

Now, construct a graph whose vertices are the points in U'. Add an edge to the graph between b and b' whenever there is a $g \in V$ with $g \in R(b) \cap R(b')$. A cycle in this graph corresponds to a cycle in R, as defined above. Therefore, the graph must be a forest (a graph with no cycles). A finite graph with no cycles and $|U'|$ nodes can have at most $|U'| - 1$ edges. This proves that $\pi_{U'}(R) \le |U'| - 1$. ∎

The following lemma is the main ingredient in the construction of the monotone factor ϕ, which is given in the next section.

LEMMA 6.2: *If ϵ is chosen sufficiently small and k is chosen to be large, then the following holds. If $(s, \mathbf{K})$ is good and $R_{s,\mathbf{K}}$ is a society which is a refinement of $Q_{s,\mathbf{K}} (= S_{\gamma_{s,\mathbf{K}}})$, then there is a society $P_{s,\mathbf{K}} \subseteq R_{s,\mathbf{K}}$ such that*

$$\begin{aligned}(13)\qquad &\mu_{s,\mathbf{K}}(\{g \in \mathcal{F}(s,\mathbf{K})\colon g \in P_{s,\mathbf{K}}(b) \text{ for more than one } b \in \mathcal{F}(s,\mathbf{K})\}) \\ &\quad \le 2^{-((h(p)-h(q))/2)\bar{l}} + \mu_{s,\mathbf{K}}(\{g\colon g \text{ is bad}\}) + \nu_{s,\mathbf{K}}(\{b\colon b \text{ is bad}\}).\end{aligned}$$

Proof: We assume here that $q_0 > 1/2$. The other case is similar. Let $P_{s,\mathbf{K}}$ be a refinement of $R_{s,\mathbf{K}}$ with $\pi_{U'}(P_{s,\mathbf{K}}) < |U'|$, which is guaranteed to exist by Lemma 6.1. Let $g \in \mathcal{F}(s,\mathbf{K})$ be such that $g \in P_{s,\mathbf{K}}(b)$ for exactly one boy $b \in U'$ and also suppose that b is good. We will show that for all $b' \ne b$, $g \notin P_{s,\mathbf{K}}(b')$.

Suppose $g \in P_{s,\mathbf{K}}(b) \cap P_{s,\mathbf{K}}(b')$. Since $b' \notin U'$, the fraction of $K \in \mathbf{K}_4$ with $b'_K \in B \setminus L_{a^*}$ is greater than ϵ and so since b is a good boy, there must exist $K \in \mathbf{K}_4$ with

$$b_K \in L_a \subset A^k \setminus B, \quad \text{so } a < a^* \quad \text{and}$$
$$b'_K \in L_{a'} \subset B \setminus L_{a^*}, \quad \text{so } a' > a^*.$$

By Lemma 4.2, $S_{\bar\gamma}(b_K) \cap S_{\bar\gamma}(b'_K) = \emptyset$. Since $R_{s,\mathbf{K}}$ is a refinement of $Q_{s,\mathbf{K}}$, there is no possible choice for g_K under $\gamma_{s,\mathbf{K}}$ and this provides a contradiction. Therefore g lies in $P_{s,\mathbf{K}}(b)$ for exactly the one good boy $b \in \bar{\mathcal{F}}(s,\mathbf{K})$.

What this shows is that if g belongs to more than one $P_{s,\mathbf{K}}(b)$ then either

- for all b with $g \in P_{s,\mathbf{K}}(b)$, b is not good, or

- there exist distinct $b, b' \in U'$ with $g \in P_{s,\mathbf{K}}(b) \cap P_{s,\mathbf{K}}(b')$. The number of such F is controlled by Lemma 6.1.

Girls g of the first type must lie in $\mathcal{F}(s,\mathbf{K}) \setminus P_{s,\mathbf{K}}(\{b : b \text{ is good}\})$. Since $P_{s,\mathbf{K}}$ is a society,

$$\mu_{s,\mathbf{K}}(\mathcal{F}(s,\mathbf{K}) \setminus P_{s,\mathbf{K}}(\{b\colon b \text{ is good}\})) \leq \nu_{s,\mathbf{K}}(\{b\colon b \text{ is not good}\}).$$

By Lemma 6.1, there are less than $|U'|$ girls g of the second type. Lemma 5.5 tells us that

$$|U'| \leq 2^{(h(q)+3\epsilon+h_{cnt}(\epsilon))l}.$$

If g belongs $P_{s,\mathbf{K}}(b)$ for more than one b and is of the second type, then either g is good, in which case $\mu_{s,\mathbf{K}}(g) \leq 2^{l(h(p)-\epsilon)}$, or else it is bad. Therefore,

$$\begin{aligned}
&\mu_{s,\mathbf{K}}(\{g \in \mathcal{F}(s,\mathbf{K})\colon g \text{ belongs to more than one } P_{s,\mathbf{K}}(b)\}) \\
&\leq 2^{\bar{l}(h(q)+3\epsilon+h_{cnt}(\epsilon))}2^{-l(h(p)-\epsilon)} + \mu_{s,\mathbf{K}}(\{g : g \text{ is bad}\}) + \nu_{s,\mathbf{K}}(\{b : b \text{ is bad}\}) \\
&\leq 2^{-\bar{l}((1-\epsilon)h(p)-h(q)-4\epsilon-h_{cnt}(\epsilon))} + \mu_s(\{g\colon g \text{ is bad}\}) + \nu_s(\{b : b \text{ is bad}\}) \\
&\leq 2^{-\bar{l}((h(p)-h(q))/2)} + \mu_s(\{g\colon g \text{ is bad}\}) + \nu_s(\{b\colon b \text{ is bad}\})
\end{aligned}$$

when ϵ is small and k is large. ∎

We end this section by constructing the set of societies (one for each pair $(s,\mathbf{K})$) which will give our finitary code. In order for our finitary code ϕ, explained briefly in Section 2, to be well defined we need to have consistency in our choice of societies with higher-order skeletons, so that the definition of ϕ does not get changed at a later stage of the construction. In order to use Lemma 6.2, we also want our societies to be **minimal**, in the sense that they contain no proper refinements.

Define a system of societies $P = (P_{s,\mathbf{K}})_{(s,\mathbf{K})}$ as follows. For each skeleton s of order one and $\mathbf{K}$ covering s, let $P_{s,\mathbf{K}}$ be any minimal refinement of $Q_{s,\mathbf{K}}$. If $P_{s,\mathbf{K}}$ has been defined whenever s is of order less than n, let s have order n and take $s_1, \ldots, s_t$ to be its maximal subskeletons and let $\mathbf{K}_t = \mathbf{K}|_{s_i}$ for $i = 1, \ldots, t$. Let $P_{s,\mathbf{K}}$ be any minimal society $\subseteq P_{s_1,\mathbf{K}_1} \times \cdots \times P_{s_t,\mathbf{K}_t}$. Note that by (9), $Q_{s,\mathbf{K}} = Q_{s_1,\mathbf{K}_1} \times \cdots \times Q_{s_t,\mathbf{K}_t}$ is itself a product society, and so $P_{s,\mathbf{K}}$ must be a refinement of $Q_{s,\mathbf{K}}$. Thus, inequality (13) applies to all $P_{s,\mathbf{K}}$ where $(s,\mathbf{K})$ is good.

7. Construction of the finitary code

We now have all of the ingredients needed to prove our main theorem, stated in Theorem 1.2.

Proof of Theorem 1.2: Fix ϵ as in Lemma 6.2 and then k as in Lemmas 5.1–5.4. Let P be the system of societies defined at the end of the last section. By Lemma 5.1, for almost every $x \in X$,

$$\lim_{t\to\infty} \mu_{s_t(x),\mathbf{K}^x|_{s_t(x)}}(\{g \in \mathcal{F}(s_t(x),\mathbf{K}^x|_{s_t(x)}) : g \text{ is not good}\}) = 0.$$

By Lemma 5.3, for almost every $x \in X$,

$$\lim_{t\to\infty} \nu_{s_t(x),\mathbf{K}^x|_{s_t(x)}}(\{b \in \bar{\mathcal{F}}(s_t(\hat{x}),\mathbf{K}^x|_{s_t(x)}) : b \text{ is not good}\}) = 0.$$

By Lemma 5.4, for a.e. $x \in X$, $(s_t(x), \mathbf{K}^x|_{s_t(x)})$ is good for all t sufficiently large. Therefore by Lemma 6.2, for almost every x,

$$\lim_{t\to\infty} \mu_{s_t(x),\mathbf{K}^x|_{s_t(x)}}(\{g \in \mathcal{F}(s_t(x),\mathbf{K}^x|_{s_t(x)}) : g \text{ belongs to more than one } P_{s_t(x),\mathbf{K}^x}(b)\}) = 0. \tag{14}$$

For any $x \in X(x')$ where (14) holds, we can define our finitary code ϕ. Write q_t and r_t for the first and last coordinates of $s_t(x)$ for all t. For any t such that the filler x_t for $s_t(x)$ lies in a unique $P_{s_t(x),\mathbf{K}^x|_{s_t(x)}}(b)$, we set

$$\phi(x)[q_t, r_t] = b.$$

The societies $P_{s,\mathbf{K}}$ in P were constructed so that if $t' > t$, then the assignments given by $P_{s_t(x),\mathbf{K}^x|_{s_t(x)}}$ are consistent with those from $P_{s_{t'}(x),\mathbf{K}^x|_{s_{t'}(x)}}$, so the map ϕ is well-defined. As $t \to \infty$, $q_t \to -\infty$ and $r_t \to \infty$ almost surely, and so $\phi(x)$ has all of its coordinates defined in this way.

Clearly, ϕ is measurable, monotone, finitary, and commutes with the shift map T (and T^{-1}). We will be done once we show that

$$\phi\mu = \nu. \tag{15}$$

It suffices to show the same equality holds on the fibers $X(x)$: to prove (15), we just have to show that

$$\phi|_{X(x)}\mu_{s(x),\mathbf{K}^x} = \nu_{s(x),\mathbf{K}^x} \quad \text{for } \mu\text{-almost every } x.$$

Consider the construction of ϕ at any stage t. Let ϕ_t be the partial map taking $g \in \mathcal{F}(s_t(x),\mathbf{K}^x|_{s_t(x)})$ to its image $\phi_t(g) = b \in \bar{\mathcal{F}}(s_t(x),\mathbf{K}^x|_{s_t(x)})$ when the image is defined. Let $\lambda_{s_t(x),\mathbf{K}^x|_{s_t(x)}}$ be a coupling which respects the society $P_{s_t(x),\mathbf{K}^x|_{s_t(x)}}$. If $g \in \phi_t^{-1}(b)$, then $g \in P_{s_t(x),\mathbf{K}^x|_{s_t(x)}}(b)$ and is in no other $P_{s_t(x)}(b')$, so $\lambda_{s_t(x),\mathbf{K}^x|_{s_t(x)}}$ must couple g *only* to b. Therefore

$$\nu_{s_t(x),\mathbf{K}^x|_{s_t(x)}}(b) \geq \lambda_{s_t(x),\mathbf{K}^x|_{s_t(x)}}((\phi_t^{-1}b) \times b) \geq \mu_{s_t(x),\mathbf{K}^x|_{s_t(x)}}(\phi_t^{-1}b),$$

so at each stage t,

$$\phi_t \mu_{s_t(x),\mathbf{K}^x|_{s_t(x)}} \preceq \nu_{s_t(x),\mathbf{K}^x|_{s_t(x)}}.$$

This implies that for almost every x,

$$\phi|_{X(x)}\mu_{s(x),\mathbf{K}^x} \preceq \nu_{s(x),\mathbf{K}^x} \implies \phi|_{X(x)}\mu_{s(x),\mathbf{K}^x} = \nu_{s(x),\mathbf{K}^x}$$

since $\phi|_{X(x)}\mu_{s(x),\mathbf{K}^x}$ and $\nu_{s(x),\mathbf{K}^x}$ are probability measures whenever (14) holds. This proves (15). ∎

We conclude this section with some remarks about the two-symbol assumption for q. The marriage lemma of Keane and Smorodinsky as stated here (Lemma 6.1) states that for any society and any collection of boys U', there is a refinement of the society such that fewer than $|U'|$ girls know more than one boy in U'. For each pair $(s,\mathbf{K})$, we constructed a society $Q_{(s,\mathbf{K})}$ and two sets of boys $U \subseteq U'$ such that

- All but a set of measure ϵ of the boys are in U for large skeletons. These are the good boys.
- $|U'| \ll \#$(good girls) for large skeletons.
- If a girl g knows a boy $b \in U$ under $Q_{s,\mathbf{K}}$, then she knows only boys from U' under $Q_{s,\mathbf{K}}$ (i.e. $Q^*_{s,\mathbf{K}}(g) \cap U \neq \emptyset \implies Q^*_{s,\mathbf{K}}(g) \subseteq U'$ where $Q^*_{s,\mathbf{K}}(g) := \{b : g \in Q_{s,\mathbf{K}}(b)\}$).

The other details of the construction are non-essential. If we could construct such societies for q with more than two symbols, then we could remove the two-symbol assumption from Theorem 1.2.

It is, however, not clear how to obtain a similar triple $(Q_{s,\mathbf{K}}, U, U')$ when q is supported on more than two symbols. The structure analogous to the sets L_a are simplices which partition the elements of $\mathcal{F}(s)$ and $\bar{\mathcal{F}}(s)$ according to how many of each symbol they have. It is not clear how to construct the analog of γ^* for these simplices, which would couple them monotonically according to the natural partial order and allow us to define sets U and U' with the properties listed above. When q has more than two symbols, the set of atypical fillers for q (H in our construction) becomes more spread around in their simplex as well; it is no longer just the set of fillers with an atypical number of 1's. These are the main obstructions to completing this proof for general probability vectors q.

8. Codes for large enough entropy gaps

In this section, we show that for any q, if $p \succcurlyeq q$ and $h(p)$ is large enough, then there is a monotone factor from $B(p)$ to $B(q)$. By requiring that p has sufficiently large entropy, we are able to simplify the proof used for Theorem 1.2 considerably and drop the two-symbol assumption. Let d_q be the number of i for which $q_i \neq 0$.

THEOREM 8.1: *If $p \succcurlyeq q$ and $h(p) > \log d_q$, then there is a monotone factor $\phi: B(p) \to B(q)$.*

In particular, this implies Theorem 1.3, since if q puts probability $1/n$ on each of n symbols, then $h(q) = \log d_q$.

Because the proof of this theorem is just a very simplified version of the proof of Theorem 1.2, we only outline the proof. One reason that this proof is simpler is that we may use π, the natural monotone coupling of $B(p)$ and $B(q)$ in place of γ.

Fix $\epsilon > 0$ to be chosen later. Let s be a skeleton of total length $\bar{l}$ and let l be the length of s without its markers. Define $\mu_s(\cdot) := \mu(\cdot \mid \mathcal{F}(s))$. We say that s is **good** if $l \geq (1-\epsilon)\bar{l}$ and that $g \in \mathcal{F}(s)$ is **good** if

$$\mu_s(g) \leq 2^{-(h(p)-\epsilon)l}.$$

LEMMA 8.2: *For any $\epsilon > 0$ and k sufficiently large, the following holds: for almost all $x \in X$, $s_t(s)$ is a good skeleton and $x_{s_t(x)}$ is a good girl in $\mathcal{F}(s_t(x))$ for all t sufficiently large.*

Proof: By definition, $\bar{l} - l = k \cdot \#(\text{markers in } s)$. For k large, it is a rare event to lie in a marker, so we can take k large enough that the proportion of coordinates of $x|_{s_t(x)}$ in markers is less than ϵ for all t large enough. For all such t, $s_t(x)$ is good.

The fact that $x_{s_t(x)}$ is a good girl for all t large enough follows directly from Lemma 5 in [2]. ∎

Since q has d_q symbols,

$$\#(\bar{\mathcal{F}}(s)) \leq 2^{(\log d_q)\bar{l}}.$$

It is this fact which will make this argument easier than the proof of Theorem 1.2. It eliminates the need to have good and bad fillers in $\bar{\mathcal{F}}(s)$ — all of them are good. Let $\nu_s(\cdot) := \pi(\mathcal{F}(s) \times \cdot \mid \mathcal{F}(s) \times \bar{\mathcal{F}}(s))$.

LEMMA 8.3: *For ϵ sufficiently small and k sufficiently large, if s is a good skeleton of total length $\bar{l}$ and if Q_s is a society from $(\bar{\mathcal{F}}(s), \mu_s)$ to $(\mathcal{F}(s), \nu_s)$, then there exists a society $P_s \subseteq Q_s$ such that*

$$\mu_s\{g \in \mathcal{F}(s) : g \text{ belongs to more than one } P_s(b)\} \\ \leq 2^{-((h-\log d_q)/2)\bar{l}} + \mu_s\{g \in \mathcal{F} : g \text{ is bad}\}.$$

Proof: We will use the Marriage Lemma 6.1 with $(U, \rho) = (\bar{\mathcal{F}}(s), \nu_s)$ and $(V, \sigma) = (\mathcal{F}(s), \mu_s)$, which implies that there is a refinement P_s of Q_s with $\pi(P_s) < 2^{(\log d_q)l}$.

Now, if g is good then $\mu_s(g) \leq 2^{-(h(p)-\epsilon)l}$, so the measure of all good g which are in P_s for more than one $\bar{F}$ is at most

$$2^{-(h-\epsilon)l} 2^{(\log d_q)\bar{l}} \leq 2^{-((h(p)-\log d_q)/2)\bar{l}}$$

when ϵ is chosen small enough. Adding in the measure of those g which are not good, we get the desired estimate. ∎

For any skeleton s, let $Q_s = S_\pi$ be the independent monotone society from $(\bar{\mathcal{F}}(s), \nu_s)$ to $(\mathcal{F}(s), \mu_s)$. For each skeleton s of order 1, let P_s be any minimal society refining Q_s. Once all P_s have been defined for s of order less than n, for s of order n let P_s be a minimal society refining $P_{s_1} \times \cdots \times P_{s_t}$, where $s_1, \ldots, s_t$ are the maximal subskeletons of s.

Define ϕ in the same way as in the proof of Theorem 1.2 and the proof proceeds in exactly the same way from this point to show that ϕ: $B(p) \to B(q)$.

References

[1] K. Ball, *Poisson processes as monotone factors of Poisson processes*, Electronic Communications in Probability, to appear.

[2] M. Keane and M. Smorodinsky, *A class of finitary codes*, Israel Journal of Mathematics **26** (1977), 352–371.

[3] M. Keane and M. Smorodinsky, *Bernoulli schemes of the same entropy are finitarily isomorphic*, Annals of Mathematics **109** (1979), 397–406.

[4] K. Petersen, *Ergodic Theory*, Cambridge University Press, Cambridge, 1983.

[5] Ja. G. Sinai, *On a weak isomorphism of transformations with invariant measure*, Doklady Akademii Nauk SSSR **147** (1962), 797–800.

[6] V. Strassen, *The existence of probability measures with given marginals*, Annals of Mathematical Statistics **36** (1965), 423–439.