*Ergod Th & Dynam Sys* (1990), **10**, 687-715
*Printed in Great Britain*

# Bernoulli shifts of the same entropy are finitarily and unilaterally isomorphic

ANDRÉS DEL JUNCO
*Department of Mathematics, University of Toronto, Toronto, M5S 1A1, Canada*

(*Received* 7 *November* 1988 *and revised October* 1989)

*Abstract* Let $p=\{p_1, \ldots, p_n\}$ and $q=\{q_1, \ldots, q_m\}$ be finite probability vectors, each having at least three non-zero components, such that $-\sum_{i=1}^{n} p_i \log p_i = -\sum_{i=1}^{m} q_i \log q_i$ Let $C=\{1, \ldots, n\}$, $D=\{1, \ldots, m\}$ and let $(C^{\mathbb{Z}}, p^{\mathbb{Z}}, \sigma)$ and $(D^{\mathbb{Z}}, q^{\mathbb{Z}}, \tau)$ be the corresponding Bernoulli shifts Then there exists an isomorphism $\phi$ between these shifts such that for a a $x \in C^{\mathbb{Z}}$ $\phi(x)(0)$ is determined by finitely many of the future co-ordinates $x(0), x(1), \ldots$ and for a a $y \in D^{\mathbb{Z}}$ $\phi^{-1}(y)(0)$ is determined by finitely many of the co-ordinates $\ldots y(-1), y(0), y(1), \ldots$

## 1 *Introduction*

Let $C=\{1, \ldots, n\}$ and $D=\{1, \ldots, m\}$ be finite sets Let $X = C^{\mathbb{Z}}$ denote the space of two-sided sequences indexed by $\mathbb{Z}$ and $Y = D^{\mathbb{Z}}$, both equipped with their Borel $\sigma$-algebras Suppose that $p$ and $q$ are probability measures on $C$ and $D$ Let $\mu = p^{\mathbb{Z}}$ and $\nu = q^{\mathbb{Z}}$ denote the corresponding product measures on $X$ and $Y$ The shift transformation $\sigma$ on $X$ is defined by $\sigma(x)(i) = x(i+1)$ $\tau$ will denote the shift on $Y$ $(X, \mu, \sigma)$ and $(Y, \nu, \tau)$ are called Bernoulli shifts

A homomorphism from $(X, \mu, \sigma)$ to $(Y, \nu, \tau)$ is a measure-preserving map $\phi$ $(X, \mu) \to (Y, \nu)$ such that $\phi \circ \sigma = \tau \circ \phi$ $\mu$-a e $\phi$ is said to be finitary if $\forall j \in D$ $\phi^{-1}\{y \in Y \; y(0) = j\}$ agrees $\mu$-a e with a countable union of finite cylinder sets in $X$ Informally, for a a $x \in X$ $\phi(x)(0)$ is determined by finitely many of the co-ordinates $\ldots, x(-1), x(0), x(1), \ldots$ and, by shift invariance, the same is true of $\phi(x)(t)$ for all $t \in \mathbb{Z}$

We will call $\phi$ forgetful if for all $j \in D$ $\phi^{-1}\{y \in Y \; y(0) = j\}$ agrees a e with a set in $\mathscr{F}^+$, the future $\sigma$-algebra in $X$, that is, the $\sigma$-algebra generated by the projections $x \mapsto x(i)$, $i = 0, 1, \ldots$ Informally, $\phi(x)(0)$ is determined by $x(0), x(1), \ldots$ We will call $\phi$ finitarily forgetful if for all $j \in D$ $\phi^{-1}\{y \in Y \; y(0) = j\}$ agrees a e with a countable union of cylinder sets each of which is in $\mathscr{F}^+$ It is an easy exercise that $\phi$ is finitarily forgetful if and only if it is finitary and forgetful

$\phi$ is called an isomorphism if it has an a e inverse $\psi$, that is $\psi\phi = id$ a e The entropy $h(p)$ of $p$ is defined by

$$h(p) = -\sum_{i=1}^{n} p_i \log p_i$$

(We use $\log_2$ exclusively ) The purpose of this paper is to prove the following result

THEOREM 1 *If $h(p)=h(q)$ and $p$ and $q$ each have at least three non-zero components then there is a finitarily forgetful isomorphism $\phi$ from $(X, \mu, \sigma)$ to $(Y, \nu, \sigma)$ whose inverse is finitary*

Except for the three-state assumption Theorem 1 is a strengthening of Keane and Smorodinsky's finitary isomorphism theorem for Bernoulli shifts [**K, S**] The present paper is a continuation of the work in [**J**] where we established the existence of a finitarily forgetful homomorphism under the hypotheses of Theorem 1 Sinai [**S**] established the existence of a forgetful homomorphism under more general hypotheses ($\mu$ any ergodic $\sigma$-invariant probability and $(Y, \nu, \tau)$ any Bernoulli shift with $h(\tau)\leq h(\sigma)$) Ornstein and Weiss [**O, W**] have given another proof of Sinai's theorem. Propp [**P**] has recently simultaneously generalized Sinai's theorem and given a proof entirely analogous to the proof of Ornstein's isomorphism theorem [**O**], using the Baire category theorem as in [**B, R**]

If one removes all finitariness requirements on $\phi$ and $\phi^{-1}$, Theorem 1 asserts the existence of a forgetful isomorphism between $\sigma$ and $\tau$, which is still a new result We will call this weaker assertion Theorem 2 Because the proof of Theorem 1 is somewhat intricate we include here a considerably simpler proof of Theorem 2 using a Baire category argument The central idea of this proof is however the same as that of Theorem 1, namely the $*$-joining as defined in [**J**] In particular we observe here that under the right assumptions the $*$-operation has a certain symmetry which escaped notice in [**J**] and leads to a proof of the associativity of the $*$-operation

Theorem 1 falls short of its intended goal in two respects One of these is the troublesome three-state restriction The other is the natural conjecture that one should be able to make $\phi^{-1}$ causal, that is $\phi^{-1}(y)(0)$ depends only on the past $\ldots, y(-1), y(0)$ (It is well known [**W**] that $\phi$ and $\psi$ cannot both be forgetful unless $q$ is a re-arrangement of $p$ ) This conjecture is based on a natural desire for symmetry, but more cogently, on the fact that the desired symmetry is quite analogous to the aforementioned symmetry of the $*$-joining (In fact the $\phi^{-1}$ we obtain is in a certain sense close to being causal ) In this connection we mention Meshalkin's construction [**M**], in case $p=(\frac{1}{2},\frac{1}{8},\frac{1}{8},\frac{1}{8},\frac{1}{8})$ and $q=(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4})$, of a finitarily forgetful isomorphism with a finitarily causal inverse

As in [**K, S**] and [**J**] the three-state assumption allows us to assume that $p_1=q_1$ As in [**K, S**] and [**J**] we use the symbol 1 as a marker and $\phi$ reproduces in $y$ each occurrence of a marker in $x$ The $\phi^{-1}$ we obtain here is marker-conditionally causal in the sense that, once all occurrences of 1's in $y$ are known, then one need only look at $\ldots, y(-1), y(0)$ to determine $\ldots, x(-1), x(0)$

The plan of this paper is as follows § 2 describes the function of markers § 3 introduces the notion of a skeleton, its rank and its filler sets This section is very similar to § 3 of [**J**] although somewhat streamlined It contains most of the intricacies which are needed to achieve finitariness Roughly speaking a skeleton $\mathcal{S}$ is a configuration of 1's (markers) and 0's (blanks) A skeleton can be filled by inserting one of the symbols from $A=\{2, \ldots, n\}$ ($B=\{2, \ldots, m\}$ in case of the $Y$ fillers) in each blank spot For each $\mathscr{S}$ this is done only on a certain pre-assigned subset $I(\mathscr{S})$, in case of $X$-fillers, or $J(\mathscr{S})$, in case of $Y$-fillers, of the blank spots Later, in § 5, we assign to each $\mathscr{S}$ a measure $\pi_{\mathscr{S}}$ on $A^{I(\mathscr{S})} \times B^{J(\mathscr{S})}$, whose marginals on $A^{I(\mathscr{S})}$ and $B^{J(\mathscr{S})}$ are $p_0^{I(\mathscr{S})}$ and $q_0^{J(\mathscr{S})}$, where $p_0$ and $q_0$ denote $p$ and $q$ conditioned on $A$ and $B$ respectively We call such a joint measure a superposition

In § 4 we describe the $*$-joining which is essentially a way of combining two superpositions $\pi_{\mathscr{S}_1}$ and $\pi_{\mathscr{S}_2}$ to get a superposition $\pi_{\mathscr{S}_1} * \pi_{\mathscr{S}_2}$ which has both $\pi_{\mathscr{S}_1}$ and $\pi_{\mathscr{S}_2}$ as marginals It turns out that, in certain circumstances, the $*$-operation is associative Proposition 4 7 captures the key feature of the $*$-joining which makes it useful for coding roughly speaking, under the right assumptions, if $\pi_{\mathscr{S}_n}, \ldots, \pi_{\mathscr{S}_0}, \pi_{\mathscr{S}_{-1}}$ are superpositions and $\pi_{\mathscr{S}_{-1}}$ is close to being a code from $A^{I(\mathscr{S}_{-1})}$ to $B^{J(\mathscr{S}_{-1})}$, in the sense that for most $x \in A^{I(\mathscr{S}_{-1})}$ $x$ is contained with respect to $\pi_{\mathscr{S}_{-1}}$ in some $y \in B^{J(\mathscr{S}_{-1})}$ (that is $\pi_{\mathscr{S}_{-1}}(x, y) = \pi_{\mathscr{S}_{-1}}(\{x\} \times B^{I(\mathscr{S}_{-1})})$), then $\pi_{\mathscr{S}_n} * \cdots * \pi_{\mathscr{S}_0} * \pi_{\mathscr{S}_{-1}}$ is close to being a code no matter how far the $\pi_{\mathscr{S}_i}$ are from coding and how large $n$ is In [**K**, **S**] coding is achieved by taking the usual product measure $\pi_{\mathscr{S}_n} \times \cdots \times \pi_{\mathscr{S}_{-1}}$ and then using a marriage lemma to perturb it in a way which respects any coding accomplished by the $\pi_{\mathscr{S}_i}$ and makes it close to coding However the perturbation no longer has the $\pi_{\mathscr{S}_i}$ as marginals The advantage of our approach is that no perturbation is required, which is what allows us to achieve forgetfulness

In § 5 we define $\pi_{\mathscr{S}}$ for each $\mathscr{S}$ For $\mathscr{S}$ of odd rank we arrange matters so that by Proposition 4 7 $\pi_{\mathscr{S}}$ is close to coding from $A^{I(\mathscr{S})}$ to $B^{J(\mathscr{S})}$ For even rank we make the coding go the other way The $\pi_{\mathscr{S}}$ are consistent in the sense that whenever $\bar{\mathscr{S}}$ is a subskeleton of $\mathscr{S}$ $\pi_{\mathscr{S}}$ has marginal $\pi_{\bar{\mathscr{S}}}$ The consistency of the $\pi_{\mathscr{S}}$ allows us to combine them to obtain a joining of the Bernoulli shifts $\sigma$ and $\tau$ and we are then able to show that this joining in fact arises from an isomorphism with the desired properties

In § 6 we give the simpler proof of the non-finitary Theorem 2 It is essentially a much less careful version of the proof of Theorem 1 If the reader wants to read only Theorem 2 he should read §§ 2, 4 and 6, which form a logically self-contained unit

## 2 *Markers*

Lemma 2 of [**K**, **S**] enables us to assume that $p(1) = q(1)$ As in [**K**, **S**] the symbol 1 will be used as a marker in $X$ and $Y$, so we review some facts from [**K**, **S**] $X$ and $Y$ are fibred by the positions of marker occurrences as follows For $x \in X$, $\hat{x} \in \hat{X} = \{0, 1\}^{\mathbf{Z}}$ is defined by

$$\hat{x}(i) = \begin{cases} 1 & \text{if } x(i) = 1 \\ 0 & \text{otherwise} \end{cases}$$

For $\xi \in \hat{X}$, $X(\xi)$ denotes the fibre over $\xi$

$$X(\xi) = \{x \in X \quad \hat{x} = \xi\}$$

The projection of $\mu$ onto $\hat{X}$, denoted by $\hat{\mu}$, is the product measure $\hat{p}^{\mathbf{Z}}$ where $\hat{p}(1) = p(1)$, $\hat{p}(0) = 1 - p(1)$ We make parallel definitions for $Y$ and evidently $\hat{\mu} = \hat{\nu}$

We denote by $\mu_\xi$ (respectively $\nu_\xi$) the conditional measure on $X(\xi)$ (respectively $Y_\xi$) so

$$\mu = \int_{\hat{X}} \mu_\xi \, d\hat{\mu}(\xi)$$

Setting $A=\{2, \dots, n\}$, $B=\{2, \dots, m\}$ and $I(\xi)=\{t \in \mathbb{Z} : \xi(t)=0\}$, $X(\xi)$ is naturally identified with $A^{I(\xi)}$ and with this identification $\mu_\xi$ is $p_0^{I(\xi)}$ where $p_0$ denotes $p$ conditioned on $A$. Similarly $\nu_\xi$ is $q_0^{I(\xi)}$ where $q_0$ is $q$ conditioned on $B$.

### 3. *Skeleta*

We denote an interval $\{i, i+1, \dots, j\}$ in $\mathbb{Z}$ by $[i, j]$, $(i-1, j]$, $[i, j+1)$ or $(i-1, j+1)$. We will be dealing with sequences $\gamma$ indexed by a subset $I$ of $\mathbb{Z}$ with entries chosen from a symbol set $\Gamma$. Thus formally $\gamma \in \Gamma^I$, that is $\gamma : I \to \Gamma$ and $I$ is the domain of $\gamma$. For $I' \subset I$ we will often denote the restriction of $\gamma$ to $I'$ by $\gamma(I')$ (rather than $\gamma | I'$).

Let $N_0 < N_1 < N_2 < \cdots$ be a sequence of positive integers to be specified later. For $r \geq 0$ by a skeleton of rank $r$ we mean the pair $(r, \mathcal{S})$ where $\mathcal{S}$ is a sequence of 0's (blanks) and 1's (markers) indexed by a finite interval $I$ in $\mathbb{Z}$ which has the form

$$0^{m_1} 1^{n_1} 0^{m_2} 1^{n_2} \cdots 0^{m_k} 1^{n_k} \tag{3.1}$$

where $m_i > 0$, $n_i > 0$ and

$$\max\{n_i : 1 \leq i < k\} < N_r \leq n_k$$

Thus any $r'$ such that $\max\{n_i : 1 \leq i < k\} < N_{r'} \leq n_k$ is a possible rank for the configuration (3.1). We distinguish between skeleta of different rank whose associated sequences are the same, even though we will usually speak loosely of the sequence $\mathcal{S}$ as a skeleton and write $r = \operatorname{rank} \mathcal{S}$. We will say $\mathcal{S}$ of the form (3.1) has maximal rank if $r = \max\{r' : N_{r'} \leq n_k\}$. We write

$$|\mathcal{S}| = \{t \in I : \mathcal{S}(t) = 0\},$$

the set of blank indices of $\mathcal{S}$. We write $l(\mathcal{S})$ for $\#|\mathcal{S}|$.

By a subskeleton $\bar{\mathcal{S}}$ of the skeleton $\mathcal{S}$ we mean the restriction of $\mathcal{S}$ to a subinterval $J$ or $I$ ending with a full marker run of $\mathcal{S}$ (i.e. $\bar{\mathcal{S}}(1+\max J$ or $\max J = \max I) = 0)$ such that $\bar{\mathcal{S}}$ is itself a skeleton with a rank not greater than that of $\mathcal{S}$. If $\bar{\mathcal{S}}$ is a subskeleton of $\mathcal{S}$ we write $\bar{\mathcal{S}} < \mathcal{S}$. For any $j \in |\mathcal{S}|$ the restriction of $\mathcal{S}$ to $I \cap [j, \infty)$ is a subskeleton of $\mathcal{S}$ with the same rank $r$ as $\mathcal{S}$. (Note that this is the only sort of restriction of $\mathcal{S}$ which may have a potential rank greater than $r$.) Moreover every subskeleton of $\mathcal{S}$ of full rank is of this form. We denote this subskeleton by ${}_j\mathcal{S}$.

A subskeleton $\bar{\mathcal{S}}$ of $\mathcal{S}$ will be called rank-maximal in $\mathcal{S}$ if it is maximal among the subskeleta of $\mathcal{S}$ with the same rank as $\bar{\mathcal{S}}$, ordered by $<$. Equivalently, the domain of $\bar{\mathcal{S}}$ cannot be extended to the left in $\mathcal{S}$ without increasing the rank of $\bar{\mathcal{S}}$, for the reason that $\bar{\mathcal{S}}$ is preceded immediately to the left by a marker run in $\mathcal{S}$ of length $l \geq N_r$, $r = \operatorname{rank} \mathcal{S}$.

Subskeleta $\bar{\mathcal{S}}_1$ and $\bar{\mathcal{S}}_2$ of $\mathcal{S}$ may have overlapping domains without one containing the other, but the following lemma asserts that this cannot happen for rank-maximal subskeleta.

LEMMA 3 1 *If $\bar{\mathcal{S}}_1$ and $\bar{\mathcal{S}}_2$ are rank-maximal subskeleta of $\mathcal{S}$ then either one is a sub-skeleton of the other or their domains are disjoint*

*Proof* If dom $\bar{\mathcal{S}}_1 \cap$ dom $\bar{\mathcal{S}}_2 \neq \varnothing$ then (reversing the roles of $\bar{\mathcal{S}}_1$ and $\bar{\mathcal{S}}_2$ if necessary) the final marker block in $\bar{\mathcal{S}}_1$ is a marker block of $\bar{\mathcal{S}}_2$ It follows that the restriction of $\mathcal{S}$ to dom $\bar{\mathcal{S}}_1 \cup$ dom $\bar{\mathcal{S}}_2$ is a subskeleton $\bar{\mathcal{S}}$ of $\mathcal{S}$ with the same rank as $\bar{\mathcal{S}}_2$ Since $\bar{\mathcal{S}}_2$ is rank maximal $\bar{\mathcal{S}}_2 = \bar{\mathcal{S}}$ so dom $\bar{\mathcal{S}}_1 \subset$ dom $\bar{\mathcal{S}}_2$ and min (dom $\bar{\mathcal{S}}_1$) = min (dom $\bar{\mathcal{S}}_2$) If max (dom $\bar{\mathcal{S}}_1$) < max (dom $\bar{\mathcal{S}}_2$) we necessarily have rank $\bar{\mathcal{S}}_1 <$ rank $\bar{\mathcal{S}}_2$ so $\bar{\mathcal{S}}_1 < \bar{\mathcal{S}}_2$ In the other case dom $\bar{\mathcal{S}}_1 =$ dom $\bar{\mathcal{S}}_2$ so $\bar{\mathcal{S}}_1 < \bar{\mathcal{S}}_2$ or $\bar{\mathcal{S}}_2 < \bar{\mathcal{S}}_1$ according to which has the greater rank □

If $\mathcal{S}$ is a skeleton of rank $r$ the occurrences of marker runs in $\mathcal{S}$ of length at least $N_{r-1}$ divide $\mathcal{S}$ into its rank-maximal subskeleta of rank $r-1$ If these subskeleta are $\mathcal{S}_t, \quad , \mathcal{S}_0$ listed in order from left to right we write

$$\mathcal{S} = \mathcal{S}_t \times \qquad \times \mathcal{S}_0$$

and refer to this as the rank decomposition of $\mathcal{S}$ (The apparent eccentricity in ordering, which will recur frequently, is in anticipation of the fact that all the constructions we shall make will have to be made from right to left in order to ensure forgetfulness ) Note that the rank decomposition may consist of $\mathcal{S}$ alone, with rank $r-1$ rather than $r$ We will make frequent use of induction on the rank of $\mathcal{S}$ and the fact that if $\bar{\mathcal{S}} < \mathcal{S}$ and rank $\bar{\mathcal{S}} <$ rank $\mathcal{S}$ then $\bar{\mathcal{S}} < \mathcal{S}_i$ for some $i$

Our next task is to define suitable subsets $I(\mathcal{S})$ and $J(\mathcal{S})$ of $|\mathcal{S}|$ so that $A^{I(\mathcal{S})}$ and $B^{J(\mathcal{S})}$ can play the role of filler sets First we define a method of truncating skeleta Let $0 < C_0 < C_1$ be a sequence of positive integers to be specified later If $\mathcal{S}$ is a 0-skeleton define

$$C(\mathcal{S}) = |_{i_0}\mathcal{S}|, \quad i_0 = \min \{ i \in |\mathcal{S}| \;\; l(_i\mathcal{S}) \le C_0 \}$$

Now suppose $C(\bar{\mathcal{S}})$ has been defined for all skeleta $\bar{\mathcal{S}}$ of rank less than $r$ and $\mathcal{S}$ has rank $r$, $\mathcal{S} = \mathcal{S}_t \times \qquad \times \mathcal{S}_0$ Let

$$i_0 = \min \{ i \in |\mathcal{S}| \;\; l(_i\mathcal{S}) \le C_r \}$$

and suppose $i_0 \in S_{\bar{t}}$, $t \ge \bar{t} \ge 0$ We define

$$C(\mathcal{S}) = |_{i_0}\mathcal{S}| \cup C(\mathcal{S}_{\bar{t}})$$

It is immediate by induction that the definition of $C(\mathcal{S})$ is forgetful in the sense that, for $i \in |\mathcal{S}|$, $C(_i\mathcal{S}) = C(\mathcal{S}) \cap [i, \infty)$ That is, in order to know how $C(\mathcal{S})$ looks to the right of $i$ one need only look at $\mathcal{S}$ to the right of $i$ It is also immediate that for rank $\mathcal{S} = r$

$$\# C(\mathcal{S}) \le C_0 + C_1 + \qquad + C_r$$

If $I$ and $J$ are subsets of $\mathbb{Z}$ we write $I < J$ if max $I <$ min $J$ The proof of the following lemma is immediate by induction on the rank of $\mathcal{S}$

LEMMA 3 2 *If $\bar{\mathcal{S}} < \mathcal{S}$ then either $C(\bar{\mathcal{S}}) < C(\mathcal{S})$ or $C(\bar{\mathcal{S}}) \subset C(\mathcal{S})$*

If $\mathscr{S}$ is a skeleton we set

$$R(\mathscr{S})=\{\max|\mathscr{S}|\}$$

and

$$C(\mathscr{S})=C(\mathscr{S})-R(\mathscr{S})$$

We now proceed to define $I(\mathscr{S})$ and $J(\mathscr{S})$ If rank $(\mathscr{S})=0$ set

$$I(\mathscr{S})=J(\mathscr{S})=\varnothing$$

Now let

$$0<m_1<m_2<m_3< \quad ,$$

and

$$0<M_2<M_4<M_6<$$

be sequences of positive integers to be specified later Suppose $\mathscr{S}=\mathscr{S}_t\times \quad \times\mathscr{S}_0\times\mathscr{S}_{-1}$ and $r=\text{rank}\,\mathscr{S}$ is odd Let us say $\mathscr{S}_i$ is initial in $\mathscr{S}$ if $0\le i<m_r$ and say $\mathscr{S}_i$ is principal in $\mathscr{S}$ if $i\ge m_r$ and $C(S_i)\subset C(\mathscr{S})$ Note that $\mathscr{S}_{-1}$ is neither initial nor principal Define

$$I(\mathscr{S})=\bigcup\{C(\mathscr{S}_i)\ \ \mathscr{S}_i\text{ principal}\}\cup\bigcup\{R(\mathscr{S}_i)\ \ \mathscr{S}_i\text{ initial}\}$$

and

$$J(\mathscr{S})=\bigcup\{C(\mathscr{S}_i)\ \ \mathscr{S}_i\text{ principal}\}$$

Since $S_{-1}$ is not used we have $J(\mathscr{S})\subset I(\mathscr{S})\subset C(\mathscr{S})$

Now suppose that $\mathscr{S}=\mathscr{S}_t\times \quad \times\mathscr{S}_0\times\mathscr{S}_{-1}$ and $r=\text{rank}\,\mathscr{S}$ is even For $0\le i\le t$ let

$$i=q(m_r+M_r)+s,\quad q\ge 0, 0\le s<M_r+m_r$$

and call $\mathscr{S}_i$ principal in $\mathscr{S}$ if $0\le s<M_r$ and $C(\mathscr{S}_i)\subset C(\mathscr{S})$ Call $\mathscr{S}_i$ auxiliary in $\mathscr{S}$ if $M_r\le s<M_r+m_r$ and $C(\mathscr{S}_i)\subset C(\mathscr{S})$ In other words we skip $\mathscr{S}_{-1}$ and then working from right to left we label the first $M_r$ $\mathscr{S}_i$'s principal, the next $m_r$ auxiliary and so on, as long as $C(\mathscr{S}_i)$ remains inside $C(\mathscr{S})$, after which we stop We will refer to any sequence

$$\beta=(\mathscr{S}_{j+M_r+m_r-1},\quad ,\mathscr{S}_{j+1},\mathscr{S}_j)$$

of $m_r$ auxiliary $\mathscr{S}_i$'s followed by $M_r$ principal $\mathscr{S}_i$ as a full block of $\mathscr{S}$ We set

$$|\beta|=\bigcup\{|\mathscr{S}_i|\ \ j\le i<j+M_r+m_r\},$$

$$J(\beta)=J(\mathscr{S})\cap|\beta|=\bigcup\{R(\mathscr{S}_i)\ \ j+M_r\le i<j+M_r+m_r\}$$
$$\cup\bigcup\{C(\mathscr{S}_i)\ \ j\le i<M_r\}$$

and

$$I(\beta)=I(\mathscr{S})\cap|\beta|$$
$$=\bigcup\{C(\mathscr{S}_i\ \ j\le i<j+M_r\}$$

Note that the forgetfulness of $C(\mathscr{S})$ implies the forgetfulness of $I(\mathscr{S})$ and $J(\mathscr{S})$ for $i\in|\mathscr{S}|$

$$I({}_i\mathscr{S})=I(\mathscr{S})\cap[i,\infty),$$
$$J({}_i\mathscr{S})=J(\mathscr{S})\cap[i,\infty)$$

LEMMA 3 3

(a) *If $\bar{\mathscr{S}}<\mathscr{S}$ then either $I(\bar{\mathscr{S}})\cup J(\bar{\mathscr{S}})\subset C(\mathscr{S})$ or*

$$I(\bar{\mathscr{S}})\cup J(\bar{\mathscr{S}})<C(\mathscr{S})$$

https://doi.org/10.1017/S014338570000585X Published online by Cambridge University Press

(b) *If $\bar{\mathscr{S}} < \mathscr{S}$ then either*

(i) $I(\bar{\mathscr{S}}) \subset I(\mathscr{S})$ *and* $J(\bar{\mathscr{S}}) \subset J(\mathscr{S})$ *or*

(ii) $I(\bar{\mathscr{S}}) \cap I(\mathscr{S}) = \varnothing$ *and* $J(\bar{\mathscr{S}}) \cap J(\mathscr{S}) = \varnothing$

*Proof*

(a) Since $I(\bar{\mathscr{S}}) \cup J(\bar{\mathscr{S}}) \subset C(\bar{\mathscr{S}})$ (a) follows from Lemma 3 2

(b) If rank $\bar{\mathscr{S}}$ = rank $\mathscr{S}$ then $\bar{\mathscr{S}} = {}_j\mathscr{S}$ for some $j \in |\mathscr{S}|$ so $I(\bar{\mathscr{S}}) = I(\mathscr{S}) \cap [j, \infty) \subset I(\mathscr{S})$ and similarly $J(\bar{\mathscr{S}}) \subset J(\mathscr{S})$ Thus we may assume rank $\bar{\mathscr{S}}$ < rank $\mathscr{S}$ so $\bar{\mathscr{S}} < \mathscr{S}_i$ for some $\mathscr{S}_i$ in the rank decomposition $\mathscr{S} = \mathscr{S}_t \times \quad \times \mathscr{S}_{-1}$ We deal first with the case when rank $\mathscr{S}$ is odd If $\mathscr{S}_i$ is initial or $i = -1$ then $I(\bar{\mathscr{S}}) \cap I(\mathscr{S}) = J(\bar{\mathscr{S}}) \cap J(\mathscr{S}) = \varnothing$ since $I(\bar{\mathscr{S}}) \cup J(\bar{\mathscr{S}}) \subset C(\mathscr{S}_i)$ If $\mathscr{S}_i$ is neither initial nor principal then certainly $I(\bar{\mathscr{S}}) \cap I(\mathscr{S}) = J(\bar{\mathscr{S}}) \cap J(\mathscr{S}) = \varnothing$ Finally if $\mathscr{S}_i$ is principal we have by (a) $I(\bar{\mathscr{S}}) \cup J(\bar{\mathscr{S}}) \subset C(\mathscr{S}_i)$ or $I(\bar{\mathscr{S}}) \cup J(\bar{\mathscr{S}}) < C(\mathscr{S}_i)$ In the first case we actually have $I(\bar{\mathscr{S}}) \subset C(\mathscr{S}_i) \subset I(\mathscr{S})$ (since $\mathscr{S}_i$ is principal) and similarly $J(\bar{\mathscr{S}}) \subset J(\mathscr{S})$ In the second case evidently $I(\bar{\mathscr{S}}) \cap I(\mathscr{S}) = J(\bar{\mathscr{S}}) \cap J(\mathscr{S}) = \varnothing$

If rank $\mathscr{S}$ is even the argument is exactly the same, replacing 'initial' by 'auxiliary' throughout □

Given a skeleton $\mathscr{S}$ we inductively define a family $D(\mathscr{S})$ of subskeleta of $\mathscr{S}$ as follows If rank $\mathscr{S} = 0$, $D(\mathscr{S}) = \varnothing$ If $\mathscr{S} = \mathscr{S}_t \times \mathscr{S}_{t-1} \times \quad \times \mathscr{S}_0$ we let

$$D(\mathscr{S}) = \{\mathscr{S}\} \cup \bigcup \{D(\mathscr{S}_i) \quad \mathscr{S}_i \text{ is not principal in } \mathscr{S}\}.$$

LEMMA 3 4

(a) *Each* $\bar{\mathscr{S}} \in D(\mathscr{S})$ *is rank-maximal in* $\mathscr{S}$

(b) *For distinct* $\mathscr{S}_1, \mathscr{S}_2 \in D(\mathscr{S})$, $I(\mathscr{S}_1) \cap I(\mathscr{S}_2) = J(\mathscr{S}_1) \cap J(\mathscr{S}_2) = \varnothing$

(c) *For each* $\bar{\mathscr{S}} < \mathscr{S}$ *there is a* $\hat{\mathscr{S}} \in D(\mathscr{S})$ *such that* $\bar{\mathscr{S}} < \hat{\mathscr{S}}$, $J(\bar{\mathscr{S}}) \subset J(\hat{\mathscr{S}})$ *and* $J(\bar{\mathscr{S}}) \subset J(\hat{\mathscr{S}})$

(d) *The operator D is forgetful for each* $i \in |\mathscr{S}|$

$$D({}_i\mathscr{S}) = \{{}_i\bar{\mathscr{S}} \quad \bar{\mathscr{S}} \in D(\mathscr{S}), i \in |\bar{\mathscr{S}}|\} \cup \{\bar{\mathscr{S}} \in D(\mathscr{S}) \quad i < \min |\bar{\mathscr{S}}|\}$$

*Proof* The proofs of (a), (b) and (d) are more or less immediate by induction To prove (c) suppose $\bar{\mathscr{S}} < \mathscr{S}$ If rank $\bar{\mathscr{S}}$ = rank $\mathscr{S}$ then $\bar{\mathscr{S}} = {}_i\mathscr{S}$ for some $i \in |\mathscr{S}|$ and we can take $\hat{\mathscr{S}} = \mathscr{S}$ If rank $\bar{\mathscr{S}}$ < rank $\mathscr{S}$ then $\bar{\mathscr{S}} < \mathscr{S}_i$ for some $\mathscr{S}_i$ in the decomposition $\mathscr{S} = \mathscr{S}_t \times \quad \times \mathscr{S}_0$ If $\mathscr{S}_i$ is principal then $I(\bar{\mathscr{S}}) \subset I(\mathscr{S})$ and $J(\bar{\mathscr{S}}) \subset J(\mathscr{S})$ so we can again take $\hat{\mathscr{S}} = \mathscr{S}$ If $\mathscr{S}_i$ is not principal then $D(\mathscr{S}_i) \subset D(\mathscr{S})$ so the result follows by induction □

## 4 The $*$-joining

In this section we will be dealing with probability measures on various finite sets It will be convenient to adopt the following notational conventions for the rest of the paper If $E$ is a finite set and $P$ is the partition of $E$ into points, for any cartesian product $X$ in which $E$ is a factor $P$ will also denote the partition of that product according to the $E$-co-ordinate Thus for $p \in P$, $p$ denotes a subset of $E$ or of $X$ depending on the context

If $\rho$ is a probability measure on $X$ and $\alpha$ is a subset of $X$, $d_\rho(P \mid \alpha)$ will denote the conditional distribution of $P$ given $\alpha$, with respect to the measure $\rho$ Thus $d_\rho(P \mid \alpha)$ is in a natural way a measure $\rho$ on $E$ If $\alpha = X$, $d_\rho(P \mid \alpha)$ is the marginal of $\rho$ on $E$ and we denote it simply by $d_\rho P$

If $F$ is another finite set and $Q$ is the partition of $F$ into points then $PQ$ will denote the partition of $E \times F$ into points, which according to our convention is the supremum of the partitions $P$ and $Q$ on $E \times F$ If $p \in P$ and $q \in Q$, $pq \in PQ$ denotes the intersection of $p$ and $q$ considered as subsets of $E \times F$ Of course all these considerations extend to any product $X$ having $E \times F$ as a factor If $\rho$ and $\sigma$ are probability measures on $E$ and $F$ by a joining of $\rho$ and $\sigma$ we mean any measure $\lambda$ on $E \times F$ whose marginals are $\rho$ and $\sigma$ We write $P \perp Q\ (\lambda)$ if $\lambda = \rho \times \sigma$

Now suppose $\sigma$ and $\rho$ are probability measures on finite sets $F$ *and* $E$ each of which it totally ordered, so that

$$F = \{f_1 < f_2 < \quad < f_s\}, \quad E = \{e_1 < e_2 < \quad < e_r\}$$

We define a joining $\sigma \quad \rho$ of $\sigma$ and $\rho$ as follows Let

$$0 = y_0 < y_1 < \quad < y_s = 1$$

be points in $[0, 1]$ such that

$$\lambda(y_{i-1}, y_i) = \sigma(f_i) \quad \text{for } 1 \leq i \leq s,$$

where $\lambda$ denotes Lebesgue measure Similarly let

$$0 = x_0 < x_1 < \quad < x_r = 1$$

be such that

$$\lambda(x_{i-1}, x_i) = \rho(e_i) \quad \text{for } 1 \leq i \leq r$$

Define a joining $\sigma \quad \rho$ of $\sigma$ and $\rho$ by

$$(\sigma \quad \rho)(f_j, e_i) = \lambda((y_{j-1}, y_j) \cap (x_{i-1}, x_i))$$

Denoting by $Q$ and $P$ the partitions of $F$ and $E$ into points, $\sigma \quad \rho$ has the useful property that, in the joining $\sigma \quad \rho$, there are at most $\#F - 1$ atoms $p \in P$ which are split by $Q$ (that is, $(\sigma \quad \rho)(pq_1) > 0$ and $(\sigma \quad \rho)(pq_2) > 0$ for distinct $q_1, q_2 \in Q$)

The proof of the following lemma is immediate from the definition of $\sigma \quad \rho$

LEMMA 4 1 *Suppose $F$, $E_2$ and $E_1$ are finite totally ordered sets and $E_2 \times E_1$ is given the lexicographic ordering* $(e_2, e_1) < (e_2', e_1') \Leftrightarrow e_2 < e_2' \vee (e_2 = e_2' \wedge e_1 < e_1')$ *Let $Q$, $P_2$ and $P_1$ denote the partitions of $F$, $E_2$ and $E_1$ into points Suppose $\sigma$ and $\rho$ are probabilities on $F$ and $E_2 \times E_1$ and let*

$$\gamma = \sigma \quad \rho, \quad \bar{\gamma} = d_\gamma(QP_2),$$

*so $\bar{\gamma}$ is the marginal of $\gamma$ on $F \times E_2$ Then*

$$\bar{\gamma} = d_\gamma(Q) \quad d_\gamma(P_2) = d_\sigma(Q) \quad d_\rho(P_2),$$

*and for $p_2 \in P_2$*

$$\begin{aligned} d_\gamma(QP_1 | p_2) &= d_\gamma(Q | p_2) \quad d_\gamma(P_1 | p_2) \\ &= d_{\bar{\gamma}}(Q | p_2) \quad d_\rho(P_1 | p_2) \end{aligned}$$

Next suppose $\pi_2$ and $\pi_1$ are probability measures on $E_2 \times F_2$ and $E_1 \times F_1$ where $F_2$ and $E_1$ are totally ordered For $i = 2, 1$ $P_i$ and $Q_i$ denote the partitions of $E_i$ and $F_i$ into points We define a joining

$$\pi = \pi_2 * \pi_1$$

of $\pi_2$ and $\pi_1$ on $(E_2 \times E_1) \times (F_2 \times F_1)$ by decreeing that

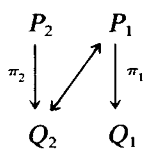
$$d_\pi(P_2) = d_{\pi_2}(P_2), \quad d_\pi(Q_1) = d_{\pi_1}(Q_1),$$
$$P_2 \perp Q_1 \quad (\pi)$$

and

$$d_\pi(Q_2 P_1 \mid p_2 q_1) = d_{\pi_2}(Q_2 \mid p_2) \quad d_{\pi_1}(P_1 \mid q_1) \quad \forall p_2 \in P_2,\ q_1 \in Q_1$$

Note that this is meaningful since we assumed $F_2$ and $E_1$ to be totally ordered

It will be important that the definition of $\pi_2 * \pi_1$ is symmetric with respect to inter-changing the roles of $E_1$ and $F_2$ and of $E_2$ and $F_1$ To be more precise, if $\pi_2'$ is the measure on $F_2 \times E_2$ corresponding to $\pi_2$ under the natural identification of $F_2 \times E_2$ with $E_2 \times F_2$, and $\pi_1'$ is the measure on $F_1 \times E_1$ corresponding to $\pi_1$ then $\pi_1' * \pi_2'$ is meaningful (since $F_2$ and $E_1$ are totally ordered) as a measure on $(F_1 \times F_2) \times (E_1 \times E_2)$ and corresponds to $\pi_2 * \pi_1$ It is useful to keep the following picture of $\pi_2 * \pi_1$ in mind




where the arrow indicates that, conditioned on any atom of $P_2 Q_1$, $Q_2$ and $P_1$ are highly correlated, while $P_2$ and $Q_1$ are independent The symmetry we have just discussed should be viewed as symmetry under rotation of the picture by 180°

LEMMA 4 2 *$\pi = \pi_2 * \pi_1$ is a joining of $\pi_2$ and $\pi_1$ Moreover with the above notation $P_2 \perp P_1 Q_1$ $(\pi)$ and $Q_1 \perp P_2 Q_2$ $(\pi)$*

*Proof* Since

$$d_\pi(Q_2 P_1 \mid p_2 q_1) = d_{\pi_2}(Q_2 \mid p_2) \quad d_{\pi_1}(P_1 \mid q_1)$$

and the dot-operation is a joining we have

$$d_\pi(Q_2 \mid p_2 q_1) = d_{\pi_2}(Q_2 \mid p_2)$$

and averaging over $q_1$

$$d_\pi(Q_2 \mid p_2) = d_{\pi_2}(Q_2 \mid p_2)$$

Since $d_\pi(P_2) = d_{\pi_2}(P_2)$ it follows that

$$d_\pi(Q_2 P_2) = d_{\pi_2}(Q_2 P_2),$$

that is $\pi$ has marginal $\pi_2$ By symmetry it also has marginal $\pi_1$ Moreover we have just observed that $Q_2$ and $Q_1$ are $P_2$-conditionally independent in the joining $\pi$ Since $Q_1$ and $P_2$ are independent it follows that $Q_2 P_2 \perp Q_1$ $(\pi)$ By symmetry we also have $P_2 \perp P_1 Q_1$ $(\pi)$ □

The next lemma asserts that under the right assumptions the $*$-operation is associative Suppose that $\pi_i$ is a probability on $E_i \times F_i$ for $i = 3, 2, 1$ and that $F_3$, $F_2$, $E_2$ and $E_1$ are totally ordered Give $F_3 \times F_2$ the reverse lexicographic ordering

$$(f_3, f_2) < (f_3', f_2') \Leftrightarrow (f_2 < f_2') \vee (f_2 = f_2' \wedge f_3 < f_3'),$$

and put the usual lexicographic ordering on $E_2 \times E_1$ Under these assumptions both $\pi_3 * (\pi_2 * \pi_1)$ and $(\pi_3 * \pi_2) * \pi_1$ are defined on $(E_3 \times E_2 \times E_1) \times (F_3 \times F_2 \times F_1)$, since the appropriate sets are totally ordered

LEMMA 4 3 *Under the above assumptions* $\pi_3 * (\pi_2 * \pi_1) = (\pi_3 * \pi_2) * \pi_1$

*Proof* As before, $P_i$ and $Q_i$ will denote the partitions of $E_i$ and $F_i$ into points We will show that $\pi = \pi_3 * (\pi_2 * \pi_1)$ has the following alternative description

$$d_\pi(P_3 P_2 Q_2 Q_1) = d_\pi(P_3) \times d_\pi(P_2 Q_2) \times d_\pi(Q_1) \tag{i}$$

$$d_\pi(Q_3 P_1 \mid p_3 p_2 q_2 q_1) = d_{\pi_3 * \pi_2}(Q_3 \mid p_3 p_2 q_2) \quad d_{\pi_2 * \pi_1}(P_1 \mid p_2 q_2 q_1), \tag{ii}$$

$$\forall p_3 \in P_3,\ p_2 \in P_2,\ q_2 \in Q_2,\ q_1 \in Q_1$$

Clearly (i) and (ii) completely determine $\pi$ The following picture of this description may be helpful

$$\begin{matrix} [P_3] & \begin{bmatrix} P_2 \\ Q_2 \end{bmatrix} & P_1 \\ Q_3 & & [Q_1] \end{matrix}$$

The boxed partitions are jointly independent of each other Note that the above picture is symmetric under the rotation by 180° Observing that rotation by 180° interchanges the reverse and usual lexicographic orderings one sees that the same description is valid for $(\pi_3 * \pi_2) * \pi_1$, establishing the lemma

To see that (i) holds observe that by Lemma 4 2 $P_3 \perp P_2 P_1 Q_2 Q_1$ $(\pi)$ and $P_2 Q_2 \perp Q_1$ $(\varepsilon)$, which implies (i)

As for (ii), by Lemmas 4 1 and 4 2 and the definition of $\pi$ for $p_3 \in P_3$, $q_2 \in Q_2$, $q_1 \in Q_1$ we have

$$\begin{aligned} d_\pi(Q_3 P_2 \mid p_3 q_2 q_1) &= d_\pi(Q_3 \mid p_3) \quad d_\pi(P_2 \mid q_2 q_1) \\ &= d_{\pi_3}(Q_3 \mid p_3) \quad d_{\pi_2 * \pi_1}(P_2 \mid q_2 q_1) \\ &= d_{\pi_3}(Q_3 \mid p_3) \quad d_{\pi_2}(P_2 \mid q_2) \end{aligned}$$

It follows that

$$d_\pi(Q_3 P_2 \mid p_3 q_2) = d_{\pi_s}(Q_3 \mid p_3) \quad d_{\pi_2}(P_2 \mid q_2)$$

Since we also have $P_3 \perp Q_2(\pi)$ we conclude that

$$d_\pi(P_3 P_2 Q_3 Q_2) = \pi_3 * \pi_2, \tag{iii}$$

that is $\pi$ has marginal $\pi_3 * \pi_2$

Moreover the above calculation shows that $Q_3 P_2$ and $Q_1$ are $P_3 Q_2$-conditionally independent in the joining $\cdot r$ Since from (i) we already know we get $P_3 Q_2 \perp Q_1$ $(\pi)$ we get

$$P_3 P_2 Q_3 Q_2 \perp Q_1\ (\pi) \tag{iv}$$

Finally by Lemma 4 1 and the definition of $\pi$ we have

$$\begin{aligned} d_\pi(Q_3 P_1 \mid p_3 p_2 q_2 q_1) &= d_\pi(Q_3 \mid p_3 p_2 q_2 q_1) \quad d_\pi(P_1 \mid p_3 p_2 q_2 q_1) \\ &= d_\pi(Q_3 \mid p_3 p_2 q_2) \quad d_\pi(P_1 \mid p_2 q_2 q_1) \text{ (by (iv) and Lemma 4 2)} \\ &= d_{\pi_3 * \pi_2}(Q_3 \mid p_3 p_2 q_2) \quad d_{\pi_2 * \pi_1}(P_1 \mid p_2 q_2 q_1) \text{ by (iii),} \end{aligned}$$

which establishes (ii) □

https://doi.org/10.1017/S014338570000585X Published online by Cambridge University Press

If $I$ and $J$ are finite subsets of $\mathbb{Z}$ we denote by $\mu_I$ and $\nu_J$ the measures $p_0^I$ and $q_0^J$ on $A^I$ and $B^J$ respectively We denote by $P^I$ the partition of $A^I$ into points, so in keeping with our conventions $P^I$ is also a partition of $A^{I'} \times B^J$ for any $J$ and $I' \supset I$ $Q^J$ will denote the partition of $B^J$ into points A probability measure $\pi$ on $A^I \times B^J$ will be called a superposition if it is a joining of $\mu_I$ and $\nu_J$ This includes the possibility that $I$ (or $J$) is empty then $\pi = \nu_J$ (or $\mu_I$)

We now fix once and for all total orderings of $A$ and $B$ We endow $A^I$ with the lexicographic ordering and $B^J$ with the reverse lexicographic ordering for $y, y' \in B^J$

$$y < y' \Leftrightarrow \exists j_0 \in J \text{ s t } y(j_0) < y'(j_0) \quad \text{and } y(j) = y'(j) \forall j > j_0,\ j \in J$$

Since any $A^I$ and $B^J$ are now totally ordered $\pi_2 * \pi_1$ is defined whenever $\pi_2$ and $\pi_1$ are superpositions Moreover, since rotation by 180° interchanges the usual and the reverse lexicographic orderings, $\pi_2 * \pi_1$ is symmetric with respect to this rotation If $I_3 < I_2 < I_1$ and $J_3 < J_2 < J_1$ are subsets of $\mathbb{Z}$ and $\pi_i$ is a superposition on $A^{I_i} \times B^{J_i}$ then Lemma 4 3 implies that $\pi_3 * (\pi_2 * \pi_1) = (\pi_3 * \pi_2) * \pi_1$ We will henceforth use this associativity without further comment and write simply $\pi_3 * \pi_2 * \pi_1$

A superposition $\pi$ on $A^I \times B^J$ will be called forgetful if for each $t \in \mathbb{Z}$

$$P^{I \cap (-\infty\ t)} \perp P^{I \cap [t\ \infty)} Q^{J \cap [t\ \infty)} \quad (\pi)$$

LEMMA 4 4 *Suppose $I_2$, $I_1$, $J_2$ and $J_1$ are finite subsets of $\mathbb{Z}$ such that $I_2 \cap I_1 = \varnothing$ and $J_2 \cap J_1 = \varnothing$ and suppose that $\pi_i$ is a superposition on $A^{I_i} \times B^{J_i}$ $(i = 2, 1)$ Then*

(a) *$\pi_2 \times \pi_1$ and $\pi_2 * \pi_1$ are superpositions on $A^{I_2 \cup I_1} \times B^{J_2 \cup J_1}$*

(b) *If $\pi_2$ and $\pi_1$ are forgetful then so is $\pi_2 \times \pi_1$*

(c) *If there are subsets $K_2 < K_1$ of $\mathbb{Z}$ such that $I_i \cup J_i \subset K_i$ $(i = 2, 1)$ and $\pi_2$ and $\pi_1$ are forgetful then so is $\pi_2 * \pi_1$*

*Proof*

(a) $\pi_2 \times \pi_1$ is obviously a superposition and $\pi_2 * \pi_1$ is a superposition by Lemma 4 2

(b) is easy

(c) Setting $\pi = \pi_2 * \pi_1$ we must show that

$$P^{(I_2 \cup I_1) \cap (-\infty\ t)} \perp P^{(I_2 \cap I_1) \cap [t, \infty)} Q^{(J_2 \cup J_1) \cap [t, \infty)} \quad (\pi), \tag{i}$$

and we may as well assume that $t \in K_2$ or $t \in K_1$ If $t \in K_1$ (i) becomes

$$P^{I_2 \cup (I_1 \cap (-\infty\ t))} \perp P^{I_1 \cap [t, \infty)} Q^{J_1 \cap [t\ \infty)} \quad (\pi),$$

which is true because of Lemma 4 2 and because $\pi_1$ is forgetful If $t \in K_2$, (i) becomes

$$P^{I_2 \cap (-\infty\ t)} \perp P^{(I_2 \cap [t, \infty)) \cup I_1} Q^{(J_2 \cap [t, \infty)) \cup J_1} \tag{ii}$$

Let $p_2' \in P^{I_2 \cap (-\infty, t)}$, $p_2^* \in P^{(I_2 \cap [t, \infty)}$ and $q_1 \in Q^{J_1}$ Because

$$d_\pi(Q^{J_2} P^{I_1} \mid p_2' p_2^* q_1) = d_{\pi_2}(Q^{J_2} \mid p_2' p_2^*) \quad d_{\pi_1}(P^{I_1} \mid q_1),$$

and because of Lemma 4 1 and the way $Q^{J_2}$ is ordered we see that

$$d_\pi(Q^{J_2 \cap [t, \infty)} P^{I_1} \mid p_2' p_2^* q_1) = d_{\pi_2}(Q^{J_2 \cap [t\ \infty)} \mid p_2' p_2^*) \quad d_{\pi_1}(P^{I_1} \mid q_1) \tag{iii}$$

Since $\pi_2$ is forgetful the distribution on the right above is independent of $p_2'$, so the same is true of the distribution on the left Thus in the joining $\pi$, $Q^{J_2 \cap [t\ \infty)} P^{I_1}$ and $P^{I_2 \cap (-\infty\ t)}$ are $P^{I_2 \cap [t\ \infty)} Q^{J_1}$-conditionally independent Since $P^{I_2 \cap (-\infty\ t)} \perp P^{I_2 \cap [t\ \infty)} Q^{J_1}$ by Lemma 4 2, we conclude that (ii) holds □

The following lemma may be viewed as an assertion of forgetfulness of the $*$-operation

LEMMA 4 5 *Suppose that* $\pi_2$ *and* $\pi_1$ *are superpositions on* $A^{I_2}\times B^{J_2}$ *and* $A^{I_1}\times B^{J_1}$ *and that there are subsets* $K_2 < K_1$ *of* $\mathbb{Z}$ *such that* $I_i \cup J_i \subset K_i$ *Fix* $t \in K_2$ *and let* $\pi_2^*$ *be the marginal of* $\pi_2$ *on* $A^{I_2\cap[t,\infty)}\times B^{J_2\cap[t\,\infty)}$ *If* $\pi_2$ *is forgetful then the marginal of* $\pi_2 * \pi_1$ *on* $A^{(I_2\cap[t,\infty))\cup I_1}\times B^{(J_2\cap[t,\infty))\cup J_1}$ *is* $\pi_2^* * \pi_1$

*Proof* Adopting the notation in the proof of Lemma 4 4(c) it follows from (iii) in that proof and the forgetfulness of $\pi_2$ that

$$d_\pi(Q^{J_2\cap[t\,\infty)}P^{I_1}\,|\,p_2^* q_1) = d_{\pi_2^*}(Q^{J_2\cap[t,\infty)}\,|\,p_2^*)\quad d_{\pi_1}(P^{I_1}|\,q_1)$$

for all $p_2^* \in P^{I_2\cap[t\,\infty)}$ and $q_1 \in Q^{J_1}$ We also have $P^{I_2\cap[t\,\infty)} \perp Q^{J_1}\,(\pi_2 * \pi_1)$ since $P^{I_2}\perp Q^{J_1}$ $(\pi_2 * \pi_1)$, so the result follows □

*Remark.* The symmetry of $\pi_2 * \pi_1$ allows one to conclude certain 'dual' statements from Lemmas 4 4 and 4 5 For example, the property dual to forgetfulness for a superposition $\pi$ is

$$P^{I\cap(-\infty\,t]}Q^{J\cap(-\infty\,t]} \perp Q^{J\cap(t\,\infty)}\quad(\pi)$$

Calling such a $\pi$ causal, Lemma 4 4(c) implies that, with the same assumptions on $I_i$ and $J_i$, if $\pi_2$ and $\pi_1$ are causal then so is $\pi_2 * \pi_1$

We say a superposition $\pi$ on $A^I \times B^J$ splits $p \in P^I$ (or the corresponding $x \in A^I$) if $\pi(pq) > 0$ and $\pi(pq') > 0$ for distinct $q, q' \in Q^J$ If this is not the case there is a $q \in Q^I$ such that $\pi(pq) = \pi(p)$ In this case we write $p \subset q(\pi)$ Analogous definitions apply to $q \in Q$

LEMMA 4 6 *Suppose* $I_n < I_{n-1} < \quad < I_{-1}$ *are finite non-empty subsets of* $\mathbb{Z}$, $J_i \subset I_i$ *for* $i = n, n-1, \quad, 0$, $\pi_i$ *is a superposition on* $A^{I_i}\times B^{J_i}$ *for* $i = n, n-1, \quad, 0$ *and let* $\pi_{-1} = \mu_{I_{-1}}$ *on* $A^{I_{-1}}$ *For* $j = n, n-1, \quad, -1$ *let*

$$l_j = \sum_{i=-1}^{j} \# I_i,$$

$$P^j = P^{I_j \cup I_{j-1}\cup \quad \cup I_{-1}},$$

$$\mu_j = \mu_{I_j \cup I_{j-1}\cup \quad \cup I_{-1}},$$

*and for* $j = n, n-1, \quad, 0$ *let*

$$Q^j = Q^{J_j \cup \quad \cup J_0},$$

$$\nu_j = \nu_{J_j \cup \quad \cup J_0}$$

*Let* $h = h(p_0) = h(q_0)$ *and fix* $\varepsilon > 0$ *If* $p \in P^j$ *then* $p = p^j p^{j-1} \quad p^{-1}$ *with* $p^i \in P^{I_i}$ *Let us call* $p$ *good if*

$$\mu_j(p) < 2^{-(h-\varepsilon)l_j}$$

Call $p$ completely good (c g ) if $p^i p^{i-1} \quad p^{-1} \in P^i$ is good for all $j \geq i \geq -1$ If $q^i \in Q^{J_i}$ for $j \geq i \geq 0$, call $q = q^j q^{j-1} \quad q^0 \in Q^j$ good if

$$\nu_j(q) > 2^{-(h-2\varepsilon)l_j}$$

and completely good if $q^i \quad q^0$ is good for all $j \geq i \geq 0$ For $n \geq j \geq -1$ let

$$\Pi_j = \pi_j * \pi_{j-1} * \quad * \pi_{-1}$$

Finally say that $p \in P^j$ is desirable if $p$ is not split by $\Pi_j$, $p \subset q\ (\pi_j)$ for a good $q \in Q^j$ and $p$ is completely good. Then setting

$$\rho_j = \mu_j(\bigcup\{p \in P^j : p \text{ is not desirable}\}),$$

for $j = n, n-1, \ldots, -1$ we have

$$\rho_j \leq \mu_j(\bigcup\{p \in P^j : p \text{ is not c.g.}\}) + \nu_j(\bigcup\{q \in Q^j : q \text{ is not c.g.}\}) + M \sum_{i=0}^{j} 2^{-\varepsilon(i+m)},$$

where $M = \max_{0 \leq i \leq n} \# B^{J_i} - 1$, $m = \# I_0$.

*Proof.* Notice that the definition of desirable is meaningful for $p \in P^{-1}$ and that we trivially have the estimate

$$\rho_{-1} < \mu \cup \{p \in P^{-1} : p \text{ is not c.g.}\}.$$

To prove the lemma by induction it suffices to assume that the desired estimate on $\rho_j$ holds for $j = n-1$ and prove that it holds for $j = n$. For $p \in P^n$ write

$$p = p^n p^* \quad \text{where } p^n \in P^{I_n},\ p^* \in P^{n-1}.$$

We obviously have

$$\rho_n \leq \rho_{n-1} + \mu_n(\bigcup \Gamma) \tag{i}$$

where

$$\Gamma = \{p^n p^* \in P^n : p^* \text{ is desirable but } p^n p^* \text{ is not}\}.$$

We claim that

$$\mu_n(\bigcup \Gamma) < M 2^{-\varepsilon(m+n)} + \mu_n(\bigcup\{\text{bad } p^n p^* \in P^n : p^* \text{ is c.g.}\}) + \nu_n(\bigcup\{\text{bad } q^n q^* \in Q^n : q^* \text{ is c.g.}\}), \tag{ii}$$

where bad means not good. To see this first observe that if $p \in \Gamma$ then $p$ belongs to one of the following sets:

$$E_1 = \Gamma \cup \{p \in P^n : p \text{ is split by } \Pi_n\}$$

$$E_2 = \Gamma \cap \{p \in P^n : p \subset q(\pi_n),\ q \text{ good}\}$$

$$E_3 = \Gamma \cap \{p \in P^n : p \subset q(\pi_n),\ q \text{ bad}\}$$

We estimate the measure of $\bigcup E_1$ by regarding it as a union of atoms of $P^n \times Q^n$ and conditioning it on sets of the form $p_1^n q^*$ where $p_1^n \in P^{I_n}$, $q^* \in Q^{n-1}$. In this argument all statements are to be interpreted modulo $\pi_n$-null sets. Since $\bigcup E_1$ is contained in the union of completely good $Q^{n-1}$-atoms, when conditioning on $p_1^n q^*$ we may assume that $q^*$ is c.g. Fixing $p_1^n q^*$ if $p^n p^* \in E_1$ then, since $p^*$ is contained in a good $q \in Q^{n-1}$, either $p^* \subset q^*$ and $p^n = p_1^n$ or $p^n p^* \cap p_1^n q^* = \varnothing$. In the first case

$$p^* \cap p_1^n q^* = p^n p^* = p_1^n p^*$$

must be split by the partition $Q^{J_n} \cap p_1^n q^*$ of $p_1^n q^*$, for otherwise $p^n p^*$ would not be split by $Q^n$. Thus to estimate $\Pi_n(\bigcup E_1 \mid p_1^n q^*)$ it suffices to estimate the

https://doi.org/10.1017/S014338570000585X Published online by Cambridge University Press

$p_1^n q^*$-conditional measure of desirable $p^* \in P^{n-1}$ such that

$$p^* \subset q^* \quad \text{and} \quad p^* \cap p_1^n q^* \quad \text{is split } Q^{J_n} \cap p_1^n q^*$$

Since

$$d_{\Pi_n}(Q^{J_n} P^n \mid p_1^n q^*) = d_{\pi_n}(Q^{J_n} \mid p_1^n) \quad d_{\Pi_{n-1}}(P^n \mid q^*),$$

there are at most $\# Q^{J_n} - 1$ such $p^*$, and the conditional measure $\Pi_n(p^* \mid p_1^n q^*)$ of such a $p^*$ is

$$\Pi_{n-1}(p^* \mid q^*) \le 2^{-(h-\varepsilon)l_{n-1}} / 2^{-(h-2\varepsilon)l_{n-1}} = 2^{-\varepsilon l_{n-1}},$$

since $p^* \subset q^*$ and both are completely good Thus

$$\Pi_n(\bigcup E_1 \mid p_1^n q^*) \le (\# Q^{J_n} - 1) 2^{-\varepsilon l_{n-1}} \le M 2^{-\varepsilon(n+m)},$$

whence also

$$\Pi_n(\bigcup E_1) = \mu_n(\bigcup E_1) \le M 2^{-\varepsilon(n+m)}$$

Now if $p \in E_2$ $p$ is bad, for otherwise $p$ would be desirable, so $\pi(\bigcup E_2)$ is less than the second term on the right of (ii) $\pi(\bigcup E_3)$ is clearly less than the third term on the right of (ii), which establishes (ii) By (i), (ii) and our induction hypothesis we have

$$\begin{aligned} \rho_n \le\ & \mu_{n-1}(\bigcup \{p \in P^{n-1} \quad p \text{ is not c g }\}) + \mu_n(\bigcup \{\text{bad } p^p p^* \in P^n \quad p^* \text{ is c g }\}) \\ & + \nu_{n-1}(\bigcup \{q \in Q^{n-1} \quad q \text{ is not c g }\}) + \nu_n(\bigcup \{\text{bad } q^n q^* \in Q^n \quad q^* \text{ is c g }\}) \\ & + M \sum_{i=0}^{n-1} 2^{-\varepsilon(i+m)} + M 2^{-\varepsilon(n+m)} \\ =\ & \mu_n(\bigcup \{p \in P^n \quad p \text{ is not c g }\}) + \nu_n(\bigcup \{q \in Q_n \quad q \text{ is not c g }\}) \\ & + M \sum_{i=0}^{n} 2^{-\varepsilon(i+m)} \end{aligned}$$

□

Lemma 4 6 is the property of the $*$-joining which is the key to the proof of Theorem 1 The following proposition articulates this property in a way which makes its applicability to coding clear

PROPOSITION 4 7 *Given $C \in \mathbb{Z}^+$ and $\eta > 0$ there exists $m \in \mathbb{Z}^+$ such that the following holds Suppose $I_n < I_{n-1} < \quad < I_0 < I_{-1}$ are finite subsets of $\mathbb{Z}$ such that*

$$\# I_i \le C \quad \text{for } i \ge 0$$

$$\# I_{-1} \ge m,$$

*and $J_i \subset I_i$ for $i = n, \quad , 0$ are such that*

$$\# J_i \le \# I_i - 1$$

*Suppose $\pi_i$ is a superposition on $A^{I_i} \times B^{J_i}$ for $i = n, \quad , 0$ and $\pi_{-1} = \mu_{I_{-1}}$ Set*

$$\pi = \pi_n * \pi_{n-1} * \quad * \pi_{-1},$$

*and $I = I_n \cup I_{n-1} \cup \quad \cup I_0 \cup I_{-1}$ Then*

$$\mu_I\{x \in A^I \quad x \text{ is not split by } \pi\} > 1 - \eta$$

*Proof* We adopt all the notation and terminology of Lemma 4 6 and also write

$$\bar{l}_j = \sum_{i=0}^{j} \# J_i$$

https://doi.org/10.1017/S014338570000585X Published online by Cambridge University Press

Since $\# I_i \leq C$ we have

$$\# J_i \leq \# I_i - 1 \leq (1 - C^{-1}) \# I_i$$

so

$$\bar{l}_j \leq (1 - C^{-1}) l_j$$

and

$$\max_i \# B^{J_i} \leq \# B^{C-1} = M$$

Now fix $\varepsilon > 0$ such that

$$h - 2\varepsilon > (1 - C^{-1})(h + \varepsilon) \tag{i}$$

By the a.e. Shannon–McMillan–Breiman theorem applied to the product measure $q_0^{\mathbb{N}}$ (in fact in this setting all that is needed is the strong law of large numbers) we can find $k$ such that

$$\forall K > k, \nu_{[1,K]}\{y \in B^{[1,K]} : \nu_{[1,\bar{l}]} y[1, \bar{l}] > 2^{-(h+\varepsilon)\bar{l}} \text{ for } k \leq \bar{l} \leq K\} > 1 - \eta \tag{ii}$$

(Recall that $y[1, \bar{l}]$ is the restriction of $y$ to $[1, \bar{l}]$.) Next choose $m$ so that

$$\min\{\nu_{[1,\bar{l}]}(y) : y \in B^{[1,\bar{l}]}, 0 \leq \bar{l} \leq k\} > 2^{-(h-2\varepsilon)m}, \tag{iii}$$

$$\forall K \geq m, \mu_{[1,K]}\{x \in A^{[1,K]} : \mu_{[1,l]} x[1, l] < 2^{-(h-\varepsilon)l} \text{ for } m \leq l \leq K\} > 1 - \eta \tag{iv}$$

and

$$M \sum_{i=m}^{\infty} 2^{-\varepsilon i} < \eta \tag{v}$$

Now if $q = q^n \cdots q^0 \in Q^n$ is not completely good (for the $\varepsilon$ we have fixed) then for some $j \geq 0$

$$\nu_j(q^j \cdots q^0) < 2^{-(h-2\varepsilon)l_j} < 2^{-(h-2\varepsilon)m}$$

(since $l_j = \# I_j + \cdots + \# I_{-1} \geq \# I_{-1} \geq m$) so, by (iii), $\bar{l}_j > k$. Moreover by (i)

$$\nu_j(q^j \cdots q^0) < 2^{-(h-2\varepsilon)l_j} < 2^{-(1-C^{-1})(h+\varepsilon)l_j} < 2^{-(h+\varepsilon)\bar{l}_j} \tag{vi}$$

By (ii) the $\nu_n$ measure of $q$'s in $Q^n$ such that (vi) occurs for some $\bar{l}_j > k$ is less than $\eta$ so we have

$$\nu_n(\bigcup\{q \in Q^n : q \text{ is not c.g.}\}) < \eta \tag{vii}$$

(iv) implies that

$$\mu_n(\bigcup\{p \in P^n : p \text{ is not c.g.}\}) < \eta \tag{viii}$$

In order to prove the lemma we may as well assume $\# I_{-1} = m$. Replacing $\eta$ by $\eta/3$ the proposition now follows from (v), (vii), (viii) and Lemma 4.6. □

PROPOSITION 4.8. *Given $C \in \mathbb{Z}^+$ and $\eta > 0$ there is an $m \in \mathbb{Z}^+$ such that the following hold. Suppose $J_{-1} < J_n < J_{n-1} < \cdots < J_0$ are finite subsets of $\mathbb{Z}$ such that*

$$\# J_i \leq C \quad i = n, n-1, \ldots, 0$$

$$\# J_{-1} \geq m$$

*and $I_i \subset J_i$, $i = n, n-1, \ldots, 0$ are such that*

$$\# I_i \leq \# J_i - 1$$

https://doi.org/10.1017/S014338570000585X Published online by Cambridge University Press

*Suppose* $\pi_i$ *is a superposition on* $A^{I_i} \times B^{J_i}$ *for* $i = n, n-1, \quad , 0$ *and* $\pi_{-1} = \nu_{J_{-1}}$ *Set* $\pi = \pi_{-1} * \pi_n * \pi_{n-1} * \quad * \pi$ *and* $J = J_{-1} \cup J_n \cup J_{n-1} \cup \quad \cup J_0$ *Then*

$$\nu_J\{y \in B^J \quad y \text{ is not split by } \pi\} > 1 - \eta$$

*Proof* This is a dual version of Proposition 4 7 which follows from the symmetry of the $*$-joining

5 *Construction of superpositions and proof of Theorem* 1

We now define for each skeleton $\mathcal{S}$ a superposition $\pi_{\mathcal{S}}$ on $A^{I(\mathcal{S})} \times B^{J(\mathcal{S})}$ Recall that if rank $\mathcal{S} = 0$ then $I(\mathcal{S}) = J(\mathcal{S}) = \varnothing$ so there is nothing to define Now suppose $\pi_{\mathcal{S}}$ has been defined when rank $\mathcal{S} < r$ and suppose $\mathcal{S} = \mathcal{S}_t \times \mathcal{S}_{t-1} \times \quad \times \mathcal{S}_{-1}$ has rank $r$ We deal first with the case of odd $r$ We assume $t \geq m_r$, as otherwise $J(\mathcal{S}) = \varnothing$ so $\pi_{\mathcal{S}}$ is simply $\mu_{I(\mathcal{S})}$ For each $i \geq 0$ such that $\mathcal{S}_{i+m_r}$ is principal let

$$I_i = C(\mathcal{S}_{i+m_r}), \bar{I}_i = I_i - \bigcup\{I(\bar{\mathcal{S}}) \quad \bar{\mathcal{S}} \in D(\mathcal{S}_{i+m_r}), I(\bar{\mathcal{S}}) \subset C(\mathcal{S}_{i+m_r})\}$$

and

$$J_i = C(\mathcal{S}_{i+m_r}), \bar{J}_i = J_i - U\{J(\bar{\mathcal{S}}) \quad \bar{\mathcal{S}} \in D(\mathcal{S}_{i+m_r}), J(\bar{\mathcal{S}}) \subset C(\mathcal{S}_{i+m_r})\}$$

Define a superposition $\pi_i$ on $A^{I_i} \times B^{J_i}$ by

$$\pi_i = \Pi\{\pi_{\bar{\mathcal{S}}} \quad \bar{\mathcal{S}} \in D(\mathcal{S}_{i+m_r})\} \times \mu_{\bar{I}_i} \times \nu_{\bar{J}_i}$$

Note that this makes sense by Lemma 3 4(b), (c) and is a superposition by Lemma 4 4(a) In particular observe that when $\mathcal{S}$ has rank 1, $\bar{I}_i = I_i$, $\bar{J}_i = J_i$, $D(\mathcal{S}_i) = \varnothing$ and $\pi_i = \mu_{I_i} \times \nu_{J_i}$ Now set

$$I_{-1} = \bigcup\{R(\mathcal{S}_i) \quad 0 \leq i < m_r\}, \qquad \pi_{-1} = \mu_{I_{-1}}$$

and define $\pi_{\mathcal{S}}$ on $A^{I(\mathcal{S})} \times B^{J(\mathcal{S})}$ by

$$\pi_{\mathcal{S}} = \pi_{\bar{t}} * \pi_{\bar{t}-1} * \quad * \pi_0 * \pi_{-1},$$

where $\bar{t}$ is the largest $i$ such that $\mathcal{S}_{i+m_r}$ is principal $\pi_{\mathcal{S}}$ is a superposition by Lemma 4 4(a) Note that $\pi_{\mathcal{S}}$ has a structure of the type assumed in Lemma 4 7 Of course if $t < m_r$ then $J(\mathcal{S}) = \varnothing$ so the conclusion of Lemma 4 7 holds vacuously

When $r = \text{rank}\,\mathcal{S}$ is even we proceed in a similar manner as follows Suppose $\mathcal{S} = \mathcal{S}_t \times \quad \times \mathcal{S}_{-1}$ For each principal $\mathcal{S}_i$ we set

$$I_i = C(\mathcal{S}_i), \bar{I}_i = I_i - \bigcup\{I(\bar{\mathcal{S}}) \quad \bar{\mathcal{S}} \in D(\mathcal{S}_i), I(\bar{\mathcal{S}}) \subset C(\mathcal{S}_i)\},$$

$$J_i = C(\mathcal{S}_i), \bar{J}_i = J_i - \bigcup\{J(\bar{\mathcal{S}}) \quad \bar{\mathcal{S}} \in D(\mathcal{S}_i), J(\bar{\mathcal{S}}) \subset C(\mathcal{S}_i)\}$$

and define a superposition $\pi_i$ on $A^{I_i} \times B^{J_i}$ by

$$\pi_i = \Pi\{\pi_{\bar{\mathcal{S}}} \quad \bar{\mathcal{S}} \in D(\mathcal{S}_{i+m_r})\} \times \mu_{\bar{I}_i} \times \nu_{\bar{J}_i}$$

For auxiliary $\mathcal{S}_i$ set $J_i = R(\mathcal{S}_i)$ and define $\pi_i = \nu_{R(\mathcal{S}_i)}$ on $B^{J_i}$ Letting $\bar{t}$ denote the largest $i$ such that $C(\mathcal{S}_i) \subset C(\mathcal{S})$ define $\pi_{\mathcal{S}}$ on $A^{I(\mathcal{S})} \times B^{J(\mathcal{S})}$ by

$$\pi_{\mathcal{S}} = \pi_{\bar{t}} * \pi_{\bar{t}-1} \quad * \pi_0$$

Note that if $\beta = (\mathcal{S}_i, \quad , \mathcal{S}_j)$ $(i = j + M_r + m_r - 1)$ is a full block of $\mathcal{S}$ then by associativity of the $*$-operation the marginal of $\pi$ on $A^{I(\beta)} \times B^{J(\beta)}$, which we will denote by $\pi_\beta$, is

$$\pi_\beta = d_\pi(P^{I(\beta)}, Q^{J(\beta)}) = (\pi_i * \pi_{i-1} * \cdots * \pi_{i-m_r+1}) * \pi_{i-m_r} * \cdots * \pi_j$$
$$= \pi_{-1} * \pi_{i-m_r} * \cdots * \pi_j$$

where

$$\pi_{-1} = \nu_{J_{-1}},\ J_{-1} = R(\mathscr{S}_i) \cup R(\mathscr{S}_{i-1}) \cup \cdots \cup R(\mathscr{S}_{i-m_r+1})$$

This $\pi_\beta$ has a structure of the type assumed in Proposition 4 8

For skeleta $\bar{\mathscr{S}}$ and $\mathscr{S}$ we will write $\bar{\mathscr{S}} \lhd \mathscr{S}$ if $\bar{\mathscr{S}} < \mathscr{S}$, $I(\bar{\mathscr{S}}) \cup J(\bar{\mathscr{S}}) \neq \varnothing$, $I(\bar{\mathscr{S}}) \subset I(\mathscr{S})$ and $J(\bar{\mathscr{S}}) \subset J(\mathscr{S})$ Lemma 3 3(b) says that if $\bar{\mathscr{S}} < \mathscr{S}$ then either $\bar{\mathscr{S}} \lhd \mathscr{S}$ or $I(\bar{\mathscr{S}}) \cap I(\mathscr{S}) = J(\bar{\mathscr{S}}) \cap J(\mathscr{S}) = \varnothing$ It is easy to see that $\bar{\mathscr{S}} \lhd \mathscr{S}$ if and only if $\bar{\mathscr{S}} = {}_j\mathscr{S}$ for some $j \in |\mathscr{S}|$ or $\bar{\mathscr{S}} < \mathscr{S}_i$ for some principal $\mathscr{S}_i$ in the rank decomposition of $\mathscr{S}$ and $I(\bar{\mathscr{S}}) \cup J(\bar{\mathscr{S}}) \subset C(\mathscr{S}_i)$

LEMMA 5 1 *The superpositions $\pi_{\mathscr{S}}$ are consistent in the sense that if $\bar{\mathscr{S}} \lhd \mathscr{S}$ then the marginal of $\pi_{\mathscr{S}}$ on $A^{I(\bar{\mathscr{S}})} \times B^{J(\bar{\mathscr{S}})}$ is $\pi_{\bar{\mathscr{S}}}$ The family $\{\pi_{\mathscr{S}}\}$ is translation invariant in the sense that if $\mathscr{S}'$ is the shift of $\mathscr{S}$ then $\pi_{\mathscr{S}'}$ is the shift of $\pi_{\mathscr{S}}$*

*Proof* First observe that each $\pi_{\mathscr{S}}$ is forgetful, as can be seen by induction on rank $\mathscr{S}$ using Lemma 4 4 To prove the consistency assertion by induction on rank $\mathscr{S}$ suppose $\bar{\mathscr{S}} \lhd \mathscr{S}$ If rank $\bar{\mathscr{S}} <$ rank $\mathscr{S}$ we must have $\bar{\mathscr{S}} < \mathscr{S}_i$ for some principal $\mathscr{S}_i$ in the rank decomposition of $\mathscr{S}$ and moreover $I(\bar{\mathscr{S}}) \cup J(\bar{\mathscr{S}}) \subset C(\mathscr{S}_i)$ By Lemma 3 4(c) $\bar{\mathscr{S}} \lhd \mathscr{S}'$ for some $\mathscr{S}' \in D(\mathscr{S}_i)$ By induction $\pi_{\mathscr{S}'}$ has marginal $\pi_{\bar{\mathscr{S}}}$ and by the definition of $\pi_{\mathscr{S}}$, $\pi_{\mathscr{S}}$ has marginal $\pi_{\mathscr{S}'}$, whence $\pi_{\mathscr{S}}$ has marginal $\pi_{\bar{\mathscr{S}}}$ as required

Thus we may assume rank $\bar{\mathscr{S}} =$ rank $\mathscr{S}$, so $\bar{\mathscr{S}} = {}_j\mathscr{S}$ for some $j \in |\mathscr{S}|$ With $j$ fixed, for any $I, J \subset \mathbb{Z}$, $\pi$ any measure on $A^I \times B^J$ and $\hat{\mathscr{S}}$ any skeleton such that $j \in |\hat{\mathscr{S}}|$ or $j < \min |\hat{\mathscr{S}}|$ let

$$I^* = I \cap [j, \infty), \quad \pi^* = d_\pi(P^{I^*}, Q^{J^*}), \quad \hat{\mathscr{S}}^* = \begin{cases} {}_j\hat{\mathscr{S}} & \text{if } j \in |\hat{\mathscr{S}}| \\ \hat{\mathscr{S}} & \text{if } j < \min |\hat{\mathscr{S}}| \end{cases}$$

What we must show is that $(\pi_{\mathscr{S}})^* = \pi_{\mathscr{S}^*}$ Let $\mathscr{S} = \mathscr{S}_t \times \cdots \times \mathscr{S}_{-1}$ and assume that rank $\mathscr{S}$ is odd for definiteness We may suppose $j \in |\mathscr{S}_{i+m_r}|$, $i \geq 0$ and $\mathscr{S}_{i+m_r}$ is principal, because if $j \in |\mathscr{S}_k|$ and $\mathscr{S}_k$ is not principal then what we are trying to show is trivial

Recall that

$$\pi_{\mathscr{S}} = \pi_{\bar{t}} * \pi_{\bar{t}-1} * \cdots * \pi_0 * \pi_{-1},$$

where we adopt all the notation introduced in the definition of $\pi_{\mathscr{S}}$ Now

$$S^* = \mathscr{S}^*_{i+m_r} \times S_{i-1+m_r} \times \cdots \times S_{-1}$$

If we set

$$\bar{D}(\mathscr{S}_{i+m_r}) = \{\hat{\mathscr{S}} \in D(\mathscr{S}_{i+m_r}) \ \ j \in |\hat{\mathscr{S}}| \text{ or } j < \min |\hat{\mathscr{S}}|\}$$

then Lemma 3 4 says that

$$D(\mathscr{S}^*_{i+m_r}) = \{\hat{\mathscr{S}}^* \ \ \hat{\mathscr{S}} \in \bar{D}(\mathscr{S}_{i+m_r})\}$$

From this and the fact that $C(\mathscr{S}^*_{i+m_r}) = (C(\mathscr{S}_{i+m_r}))^*$ one sees that

$$\pi_{\mathscr{S}^*} = \pi'_i * \pi_{i-1} * \cdots * \pi_0 * \pi_{-1}$$

where

$$\pi'_i = \Pi\{(\pi_{\hat{\mathscr{S}}^*} \ \ \hat{\mathscr{S}} \in \bar{D}(\mathscr{S}_{i+m_r})\} \times \mu_{\bar{I}^*_i} \times \nu_{\bar{J}^*_i}$$

Now by our induction hypothesis $\pi_{\hat{\mathscr{S}}^*} = (\pi_{\hat{\mathscr{S}}})^*$ for $\hat{\mathscr{S}} \in \bar{D}(\mathscr{S}_{i+m_r})$ so

$$\Pi\{\pi_{\hat{\mathscr{S}}^*}\ \ \hat{\mathscr{S}} \in \bar{D}(\mathscr{S}_{i+m_r})\} = \Pi\{(\pi_{\hat{\mathscr{S}}})^*\ \ \hat{\mathscr{S}} \in \bar{D}(\mathscr{S}_{i+m_r})\}$$
$$= (\Pi\{\pi_{\hat{\mathscr{S}}}\ \ \hat{\mathscr{S}} \in D(\mathscr{S}_{i+m_r})\})^*$$

and evidently

$$\mu_{I_i^*} = (\mu_{I_i})^*, \qquad \nu_{J_i^*} = (\nu_{J_i}^*)$$

Thus $\pi_i' = \pi_i^*$ Now $\pi_{\mathscr{S}}$ has marginal

$$\pi_i * \pi_{i-1} * \qquad \pi_0 * \pi_{-1}$$

and since $\pi_i$ is forgetful (Lemma 4 4(b) and the fact that all $\pi_{\hat{\mathscr{S}}}$ are forgetful) by Lemma 4 5 this measure in turn has marginal

$$\pi_i^* * \pi_{i-1} * \quad \cdot \pi_{-1} = \pi_{\mathscr{S}^*}$$

Thus $\pi_{\mathscr{S}}$ has marginal $\pi_{\mathscr{S}^*}$, which is what we wanted to prove

When $r$ is even the argument is essentially the same

Finally, the assertion about translation invariance is clear □

Our next task is to construct for a a $\xi \in \hat{X}$ a superposition $\pi_\xi$ on $A^{I(\xi)} \times B^{I(\xi)}$ (that is $\pi_\xi$ is a joining of $\mu_\xi$ and $\nu_\xi$) whose marginal on $A^{I(\mathscr{S})} \times B^{J(\mathscr{S})}$ is $\pi_{\mathscr{S}}$ for every skeleton $\mathscr{S}$ occurring in $\xi$ (the precise meaning of this will be explained below)

For $\xi \in \hat{X}$, $\mathscr{S}$ a skeleton and $I = [i, j]$ an interval in $\mathbb{Z}$ we will say $\mathscr{S}$ occurs in $\xi$ on $I$ if $I = \operatorname{dom} \mathscr{S}$, the restriction $\xi(I) = \mathscr{S}$ and $\xi(j+1) = 0$ We will say $\mathscr{S}$ is rank maximal in $\xi$ if no leftwards extension in $\xi$ of the domain of $\mathscr{S}$ yields a skeleton with the same rank as $\mathscr{S}$

Now suppose $\xi \in \hat{X}$ and each finite sequence of 0's and 1's occurs infinitely often as a block in $\xi$ This is true for $\hat{\mu}$-a a $\xi$ Given such a $\xi$ and $r > 0$ there exists a unique maximal interval $I$ in $\mathbb{Z}$ containg 0 such that $\xi(I)$ is an $r$-skeleton $I = [i_0, j_0]$, where

$$i_0 = \max\{i \in \mathbb{Z}\ \ i \le 0, \xi[i - N - r, i] = 1^{N_r}0\}$$
$$j_0 = \min\{j \in \mathbb{Z}\ \ j \ge 0, \xi[j - N_r + 1, j+1] = 1^{N_r}0\}$$

We denote the $r$-skeleton $\xi(I)$ by $\mathscr{S}_r(\xi)$ Then $\mathscr{S}_r(\xi)$ occurs in $\xi$ on $I = \operatorname{dom} \mathscr{S}(\xi)$ and is rank-maximal in $\xi$ Evidently

$$\mathscr{S}_0(\xi) < \mathscr{S}_1(\xi) < \mathscr{S}_2(\xi) < \qquad ,$$

$\mathscr{S}_r(\xi)$ is rank-maximal in $\mathscr{S}_{r-1}(\xi)$ and

$$|\mathscr{S}_r(\xi)| \nearrow I(\xi)$$

We will also write

$$I_r(\xi) = I(\mathscr{S}_r(\xi)) \quad J_r(\xi) = J(\mathscr{S}_r(\xi))$$

Having established this notation let us now fix a sequence $\eta_r > 0$ such that $\sum \eta_r < \infty$ Fix an odd $r$ and suppose that $m_i$, $M_i$, $N_i$ and $C_i$ have been chosen for all $i < r$ (only for even $i$ in case of $M_i$ ) For any $r$-skeleton $\mathscr{S} = \mathscr{S}_i \times \quad \times \mathscr{S}_1$ the sets $C(\mathscr{S}_i)$

https://doi.org/10.1017/S014338570000585X Published online by Cambridge University Press

are all bounded in size by $C_0 + \quad + C_{r-1}$, so by Lemma 4 7 we can choose $m_r$ so that for all skeleta $\mathcal{S}$ of rank $r$

$$\mu_{I(\mathcal{S})}\{x \in A^{I(\mathcal{S})} \quad \pi_{\mathcal{S}} \text{ splits } x\} < \eta_r \tag{5 1}$$

Next we can choose $N_r$ and then $C_r$ so that

$$\hat{\mu}\{\xi \in \hat{X} \quad C(\mathcal{S}_{r-1}(\xi)) \subset I_r(\xi), C(\mathcal{S}_{r-1}(\xi)) \subset J_r(\xi),$$
$$\mathcal{S}_{r-1}(\xi) \text{ is principal in } \mathcal{S}_r(\xi), \quad C(\mathcal{S}_r(\xi)) = |\mathcal{S}_r(\xi)|\} > 1 - \eta_r \tag{5 2}$$

(Note that in fact the first two conditions in the definition of the above set are redundant ) To see that this is possible set

$$G = \{\xi \in \hat{X} \quad \xi[0, N_{r-1}+2] = 01^{N_{r-1}}0\}$$

and choose $K$ so large that the $\hat{\mu}$-measure of

$$H_1 = \left\{\xi \in \hat{X} \quad \sum_{i=0}^{k} 1_G((\hat{\sigma})^i(\xi)) \geq m_r + 2\right\}$$

is greater than $1 - \eta_r/3$ ($\hat{\sigma}$ denotes the shift on $\hat{X}$) Then setting

$$F = \{\xi \in \hat{X} \quad \xi[0, N_r - 1] = 1^{N_r}\}$$

$N_r$ can be chosen so large that the $\hat{\mu}$-measure of

$$H_2 = \{\xi \in \hat{X} \quad \hat{\sigma}^i(\xi) \notin F \text{ for } 0 \leq i \leq K\}$$

is greater than $1 - \eta_r/3$ Finally, once $N_r$ has been chosen it is clear that $C_r$ can be chosen so large that the $\hat{\mu}$-measure of

$$H_3 = \{\xi \in \hat{X} \quad C(\mathcal{S}_r(\xi)) = |\mathcal{S}_r(\xi)|\}$$

is greater than $1 - \eta_r/3$ Now if $\xi \in H_1 \cap H_2$ then $\mathcal{S}_{r-1}(\xi)$ is not initial in $\mathcal{S}_r(\xi)$ because there are at least $m_r + 1$ rank maximal $r-1$ skeleta in $\mathcal{S}_r(\xi)$ to the right of $\mathcal{S}_{r-1}(\xi)$ If in addition $\xi \in H_3$ then all but the initial skeleta in the rank decomposition of $\mathcal{S}_r(\xi)$ are principal, and in particular $C(\mathcal{S}_{r-1}(\xi)) \subset I_r(\xi)$ and $C(\mathcal{S}_{r-1}(\xi)) \subset J_r(\xi)$ Since $\mu(H_1 \cap H_2 \cap H_3) > 1 - \eta_r$ we get 5 2

Now suppose that $r$ is even and $m_i$, $M_i$, $N_i$ and $C_i$ have been chosen for all $i < r$ By Lemma 4 8 $m_r$ can be chosen so that for each $r$-skeleton $S$ and full block $\beta$ of $\mathcal{S}$

$$\nu_{J(\beta)}\{y \in B^{J(\beta)} \quad \pi_\beta \text{ splits } y\} < \eta_r \tag{5 3}$$

(Recall the remarks following the definition of $\pi_{\mathcal{S}}$ for even $\mathcal{S}$)

Next we can choose $M_r$, $N_r$ and $C_r$, in that order, so that

$$\hat{\mu}\{\xi \in \hat{X} \quad C(\mathcal{S}_{r-1}(\xi))\} \subset J_r(\xi), C(\mathcal{S}_{r-1}(\xi)) \subset I_r(\xi),$$
$$\mathcal{S}_{r-1}(\xi) \text{ principal in } \mathcal{S}_r(\xi), \mathcal{S}_{r-1}(\xi) \text{ belongs to}$$
$$\text{a full block of } \mathcal{S}_r(\xi), C(\mathcal{S}_r(\xi)) = |\mathcal{S}_r(\xi)|\} > 1 - \eta_r \tag{5 4}$$

(Again the first two of the above conditions are redundant )

This is accomplished in much the same way as for odd $r$ First we choose $M_r$ much larger then $m_r$ and then we choose $N_r$ so that with high $\hat{\mu}$-probability the rank decomposition of $\mathcal{S}_r(\xi)$ contains $t$ $r-1$-skeleta with $t > L(M_r + m_r)$ and $L$ is large In particular, with high probability $\mathcal{S}_{r-1}(\xi)$ will not be among the leftmost $M_r$ $r-1$-skeleta in the rank decomposition of $\mathcal{S}_r(\xi)$ Finally $C_r$ is chosen so large that with high probability $\mathcal{S}_{r-1}(\xi)$ is principal in $\mathcal{S}_r(\xi)$ and also not among the leftmost $M_r$ $r-1$-skeleta in $\mathcal{S}_r(\xi)$, which two conditions together ensure that $\mathcal{S}_{r-1}(\xi)$ belongs to a full block of $\mathcal{S}_r(\xi)$ The remaining conditions in (5 4) are immediate

We now assume that $m_r$, $M_r$, $N_r$ and $C_r$ have been chosen for all $r$ so that (5 1) and (5 2) hold for all odd $r$ and (5 3) and (5 4) hold for all even $r$ By (5 2), (5 4) and the Borel–Cantelli lemma there is a set $\hat{X}'\subset\hat{X}$ such that $\hat{\mu}(\hat{X}')=1$ and for each $\xi\in\hat{X}'$ there is an $r_0(\xi)$ such that for $r\geq r_0(\xi)$ $\xi$ belongs to the appropriate set in (5 2) or (5 4) according as $r$ is odd or even Thus if $\xi\in\hat{X}'$ then for sufficiently large odd $r$

$$|\mathcal{S}_{r-1}(\xi)|-R(\mathcal{S}_{r-1}(\xi))=C(\mathcal{S}_{r-1}(\xi))\subset J_r(\xi)\subset I_r(\xi)$$

Since $|\mathcal{S}_{r-1}(\xi)|\nearrow I(\xi)$ we conclude

$$\bigcup_r J_r(\xi)=\bigcup_r I_r(\xi)=I(\xi)$$

Moreover for sufficiently large $r$, even or odd, we have

$$I_{r-1}(\xi)\subset C(\mathcal{S}_{r-1}(\xi))\subset I_r(\xi),$$

so the sequence $\{I_r(\xi)\}$ is eventually increasing, and similarly the same goes for $\{J_r(\xi)\}$ Thus for sufficiently large $r$ we have $\mathcal{S}_{r-1}(\xi)\lhd\mathcal{S}_r$ Whenever this is the case Lemma 5 1 implies that $\pi_{\mathcal{S}_{r-1}(\xi)}$ and $\pi_{\mathcal{S}_r(\xi)}$ are consistent measures and when it is not the case Lemma 3 3(b) says that there is no conflict between $\pi_{\mathcal{S}_{r-1}(\xi)}$ and $\pi_{\mathcal{S}_r(\xi)}$

In view of these remarks for $\xi\in\hat{X}^*$ we can define $\pi_\xi$ to be the probability measure on $A^{I(\xi)}\times B^{I(\xi)}$ whose marginal on $A^{I_r(\xi)}\times B^{J_r(\xi)}$ is $\pi_{\mathcal{S}_r}$ for each $\xi$ It is clear that $\pi_\xi$ is a joining of $\mu_\xi$ and $\nu_\xi$ and that $\pi_\xi$ has marginal $\pi_{\mathcal{S}}$ for any skeleton $\mathcal{S}$ occurring in $\xi$, since $\mathcal{S}<\mathcal{S}_r(\xi)$ for sufficiently large $r$ Lemma 5 1 implies that the family $\{\pi_\xi\}$ is shift-invariant

$$(\sigma\times\tau)(\pi_\xi)=\pi_{\hat{\sigma}(\xi)}$$

(Recall that $A^{I(\xi)}\times B^{I(\xi)}\sim X(\xi)\times Y(\xi)$, $\hat{\sigma}$ denotes the shift on $\hat{X}$ and $\sigma$ and $\tau$ the shifts on $X$ and $Y$, so $\sigma$ $X(\xi)\to X(\hat{\sigma}(\xi))$ ) Now define $\pi$ on $X\times Y$ by

$$\pi=\int_{\hat{X}}\pi_\xi\,d\hat{\mu}(\xi)$$

(We leave it to the reader to formulate and verify the measurability which makes this meaningful ) $\pi$ is a joining of $\mu$ and $\nu$ and is invariant under $\sigma\times\tau$ The proof of Theorem 1 is now completed by the following proposition

PROPOSITION 5 2 *There exists a finitarily forgetful homomorphism $\phi$ $X\to Y$ with a finitary inverse $\psi$ such that for $B\subset X\times Y$*

$$\pi(B)=\mu\{x\in X\;\;(x,\phi(x))\in B\}$$

*Proof* For $\xi\in\hat{X}'$, define

$$\phi_\xi\;\;X(\xi)\to Y(\xi)$$

https://doi.org/10.1017/S014338570000585X Published online by Cambridge University Press

by requiring that, for $x \in X(\xi)$, the restriction $\phi_\xi(x)(J_r(\xi))$ be $y_r$, whenever $x(I_r(\xi)) \subset y_r$ $(\pi_{\mathcal{S}_r(\xi)})$

To check that this definition is unambiguous suppose that $r < r'$,

$$x(I_r(\xi)) \subset y_r \quad (\pi_{\mathcal{S}_r(\xi)}),$$

$$x(I_{r'}(\xi)) \subset y_r \quad (\pi_{\mathcal{S}_r}(\xi)),$$

and

$$S_r(\xi) \lhd S_r(\xi)$$

(If the last condition does not hold then $I_r(\xi) \cap I_r(\xi) = J_r(\xi) \cap J_r(\xi) = \varnothing$, so there is no conflict between $y_r$ and $y_r$ ) Then the marginal of $\pi_{\mathcal{S}_r(\xi)}$ on $A^{I_r(\xi)} \times B^{J_r(\xi)}$ is $\pi_{\mathcal{I}_r(\xi)}$ so, regarding the various finite sequences as cylinders in $A^{I_r(\xi)} \times B^{J_r(\xi)}$, up to $\pi_{\mathcal{S}_r(\xi)}$-null sets we have

$$x(I_{r'}(\xi)) \subset y_r \subset y_{r'}(J_r(\xi))$$

and

$$x(I_r(\xi)) \subset x(I_r(\xi)) \subset y_r$$

Thus

$$\nu_{J_r(\xi)}(y_{r'}(J_r(\xi)) \cap y_r) = \pi_{\mathcal{S}_r(\xi)}(y_r(J_r(\xi)) \cap y_r)$$

$$\geq \pi_{\mathcal{S}_r} x(I_r(\xi)) = \mu_{I_r(\xi)} x(I_{r'}(\xi)) > 0$$

This means that, as cylinders in $B^{J_r(\xi)}$, $y_r(J_r(\xi)) = y_r$ (since otherwise they are disjoint), which is what we wanted to check

Next we must check that, given $\xi$, for $\mu_\xi$-a a $x \in X(\xi)$ the sequence $\phi_\xi(x)$ is defined on all of $I(\xi)$ By 5 1, for odd $r$

$$\mu_{I_r(\xi)}\{x \in A^{I_r(\xi)} \quad \pi_{S_r(\xi)} \text{ splits } x\} < \eta_r$$

By Borel–Cantelli it follows that for $\mu_\xi$-a a $x \in X(\xi)$, $x(I_r(\xi))$ is split by $\pi_{S_r(\xi)}$ for only finitely many odd $r$ Thus for sufficiently large odd $r$ $\phi_\xi(x)$ is defined on $J_r(\xi)$ and $\{J_r(\xi) \quad r \text{ odd}\}$ is an eventually increasing sequence whose union is $I(\xi)$ This means that $\phi_\xi(x)$ is defined on $I(\xi)$ This completes the proof that the definition of $\phi_\xi$ is meaningful

The shift-invariance of $\{\pi_\xi \quad \xi \in \hat{X}\}$ evidently implies that the family $\{\phi_\xi \quad \xi \in \hat{X}\}$ is shift-invariant

$$\phi_{\hat{\sigma}(\xi)} \circ \sigma = \tau \circ \phi_\xi$$

We now define $\phi \quad X \to Y$ by

$$\phi(x) = \phi_\xi(x) \quad \text{where } \xi = \hat{x}, \text{ i e } x \in X(\xi)$$

The shift invariance of $\{\phi_\xi\}$ implies that $\phi \circ \sigma = \tau \circ \phi$

In order to show that $\phi$ is finitarily forgetful we will define a one-sided version of $\phi$ We first introduce some one-sided notation Let

$$X^* = C^{[0,\infty)}, \quad Y^* = D^{[0\ \infty)}, \quad \hat{X}^* = \{0, 1\}^{[0\ \infty)}$$

$x \mapsto \hat{x}$ denotes the natural projection $X^* \to \hat{X}^*$ and the same for $Y^* \to \hat{X}$ For $\xi \in \hat{X}^*$ let

$$I^*(\xi) = \{i \quad \xi(i) = 0\} \subset [0, \infty)$$

Then

$$X^*(\xi)=\{x\in X^* \ \hat{x}=\xi\}\sim A^{I^*(\xi)}$$

and

$$Y^*(\xi)=\{y\in Y^* \ \hat{y}=\xi\}\sim B^{I^*(\xi)}$$

$x\mapsto x^*$ will denote the projection $X\to X^*$ and more generally also the projection from any two-sided sequence space to its one-sided version Thus for example if $\xi\in\hat{X}$ and $x\in X(\xi)$ then $\xi^*\in\hat{X}^*$ and $x^*\in X^*(\xi^*)$ $\mu^*$, $\nu^*$, $\hat{\mu}^*$, $\mu_\xi^*$ and $\nu_\xi^*$ $(\xi\in\hat{X}^*)$ denote the measures on $X^*$, $Y^*$, $\hat{X}^*$, $X^*(\xi)$ and $Y^*(\xi)$ naturally corresponding to $\mu$, $\nu$, $\hat{\mu}$, $\mu_\xi$ and $\nu_\xi$

For $\xi\in\hat{X}^*$ let $i_0(\xi)$ denote the least $i$ such that $\xi(i)=0$ As in the two-sided situation there exists for each $r$ a unique interval $I=[i_0(\xi), j]$ beginning with $i_0(\xi)$ such that $\xi(j+1)=0$ and $\xi(I)$ is an $r$-skeleton We denote this $r$-skeleton by $\mathscr{S}_r^*(\xi)$ and also write $I_r^*(\xi)=I(\mathscr{S}_r^*(\xi))$ and $J_r^*(\xi)=J(\mathscr{S}_r^*(\xi))$ For $\xi\in\hat{X}^*$ we define

$$\phi_\xi^* \ X^*(\xi)\to Y^*(\xi)$$

by requiring that, for $x\in X^*(\xi)\sim A^{I^*(\xi)}$, $\phi_\xi^*(J_r^*(\xi))$ be $y_r$ whenever

$$x(I_r^*(\xi))\subset y_r(\pi_{\mathscr{S}_r^*(\xi)})$$

One argues that this defines the sequence $\phi_\xi^*(x)\in Y^*(\xi)\sim B^{I^*(\xi)}$ unambiguously and on all of $I^*(\xi)$, just as we did in the two-sided situation

Moreover for $\xi\in\hat{X}$ the mappings and $\phi_\xi$ and $\phi_{\xi^*}^*$ are consistent in the sense that for $x\in X(\xi)$

$$(\phi_\xi(x))^*=\phi_{\xi^*}(x^*) \tag{i}$$

To see this just observe that for $\xi\in\hat{X}$ we must have

$$\mathscr{S}_r^*(\xi^*)<\mathscr{S}_r(\xi)$$

for all sufficiently large $r$ Since the skeleta have the same rank, $\mathscr{S}_r^*(\xi^*)\lhd\mathscr{S}_r(\xi)$ so the superpositions $\pi_{\mathscr{S}_r^*(\xi^*)}$ and $\pi_{\mathscr{S}_r(\xi)}$ are consistent This in turn allows us to argue just as in the two-sided situation that whenever

$$x(I_r^*(\xi^*))\subset y_r^* \quad (\pi_{\mathscr{S}_r^*(\xi^*)})$$

and

$$x(I_r(\xi))\subset y_r \quad (\pi_{\mathscr{S}_r(\xi)})$$

then $y_r^*$ is the restriction of $y_r$, which gives (i)

Now define $\phi^*$ $X^*\to Y^*$ by

$$\phi^*(x)=\phi_\xi^*(x) \quad \text{for } x\in X^*(\xi)$$

$\phi^*$ is a finitary mapping since, for sufficiently large $r$, $\phi^*(x)(J_r(\xi))$ (and hence $\phi^*(x)(0)$) is determined by $\pi_{\mathscr{S}_r^*(\xi)}$ and $x(I_r^*(\xi))$, both of which are determined by the restriction of $x$ to the domain of $\mathscr{S}_r(\xi)$ (i) implies that

$$(\phi(x))^*=\phi^*(x^*),$$

which in view of the finitariness of $\phi^*$ means that $\phi$ is finitarily forgetful

Next we show that for $B\subset X\times Y$

$$\pi(B)=\mu\{x \ (x,\phi(x))\in B\} \tag{ii}$$

Fix $\xi \in \hat{X}$ and suppose that $p_0 \in P^{I_{r_0}(\xi)}$ and $q_0 \in Q^{J_{r_0}(\xi)}$ We recall the convention that $p_0$ and $q_0$ can also be regarded as subsets of larger Cartesian products By the definition of $\phi$ we have

$$\begin{aligned}
&\mu_\xi\{x \in X(\xi) \ \ (x, \phi_\xi(x)) \in p_0 q_0\} \\
&\quad = \lim_{r\to\infty} \mu_\xi(\bigcup \{p \in P^{I_r(\xi)} \ \ p \subset p_0, \exists q \in Q^{J_r(\xi)} \text{ s t } p \subset q \subset q_0 \ (\pi_{\mathscr{S}_r(\xi)})\}) \\
&\quad \leqslant \lim_{r\to\infty} \mu_\xi(\bigcup \{p \in P^{I_r(\xi)} \ \ p \subset p_0, \ p \subset q_0 \ (\pi_{\mathscr{S}_r(\xi)})\}) \\
&\quad \leq \lim_{r\to\infty} \pi_{\mathscr{S}_r(\xi)}(p_0 q_0) = \lim_{r\to\infty} \pi_{\mathscr{S}_{r_0}(\xi)}(p_0 q_0) = \pi_\xi(p_0 q_0)
\end{aligned}$$

Thus whenever $B$ is a cylinder in $X(\xi) \times Y(\xi)$, and hence for all Borel $B \subset X(\xi) \times Y(\xi)$, we have

$$\mu_\xi\{x \in X(\xi) \ \ (x, \phi_\xi(x)) \in B\} \leq \pi_\xi(B)$$

Now as a function of $B$ the left hand side of the above inequality is a probability measure on $X(\xi)$ Since $\pi_\xi$ is also a probability we can replace the inequality by equality Integrating, we obtain (ii) Note that one consequence of (ii) is that $\phi$ is measure-preserving

It now remains only to construct a finitary $\psi \ \ Y \to X$ inverse to $\phi$ For $\xi \in \hat{X}'$ and even $r$ if $\mathscr{S}_{r-1}(\xi)$ lies in a full block in $\mathscr{S}_r(\xi)$ we denote this full block by $\beta_r(\xi)$ By (5 4) and the Borel-Cantelli lemma, for $\hat{\mu}$-a a $\xi$, $\beta_r(\xi)$ is defined once $r$ is sufficiently large and even Whenever $r < r'$ and $\beta_r(\xi)$ and $\beta_{r'}(\xi)$ are defined we have $|\beta_r(\xi)| \subset |\beta_{r'}(\xi)|$ and either

$$\mathscr{S}_r(\xi) \lhd \mathscr{S}_{r'}(\xi) \tag{iii}$$

or

$$I_r(\xi) \cap I_{r'}(\xi) = J_r(\xi) \cap J_{r'}(\xi) = \varnothing \tag{iv}$$

In case (iii) holds

$$I(\beta_r(\xi)) = I_r(\xi) \cap |\beta_r(\xi)| \subset I_{r'}(\xi) \cap |\beta_{r'}(\xi)| = I(\beta_{r'}(\xi))$$

and similarly $J(\beta_r(\xi)) \subset J(\beta_{r'}(\xi))$ Thus in case (iii), $\pi_{\beta_{r'}(\xi)}$, a measure on $A^{I(\beta_{r'}(\xi))} \times B^{J(\beta_{r'}(\xi))}$, has marginal $\pi_{\beta_r(\xi)}$, since each is a marginal of $\pi_\xi$ In case (iv) we have

$$I(\beta_r(\xi)) \cap I(\beta_{r'}(\xi)) = J(\beta_r(\xi)) \cap J(\beta_{r'}(\xi)) = \varnothing$$

We have observed earlier that for a a $\xi$ (iii) holds once $r$ is sufficiently large Moreover if $S_{r-1}(\xi)$ is principal in $S_r(\xi)$, which is the case for sufficiently large $r$ by (5 4), we will have

$$I(\beta_r(\xi)) \supset C(S_{r-1}(\xi))$$

Also if $r$ is sufficiently large 5 4 implies

$$|S_{r-1}(\xi)| - R(S_{r-1}(\xi)) = C(S_{r-1}(\xi)),$$

and since $\bigcup_r |S_{r-1}(\xi)| = I(\xi)$ we conclude $\bigcup_r I(\beta_r(\xi)) = I(\xi)$

We now define

$$\psi_\xi \ \ Y(\xi) \to X(\xi)$$

by requiring that $\psi_\xi(y)(I(\beta_r(\xi))) = x_r$ whenever $y(J(\beta_r(\xi))) \subset x_r$ $(\pi_{\beta_r(\xi)})$ (5 3) together with the remarks in the previous paragraph allow us to conclude, as we did for $\phi_\xi$, that for $\nu_\xi$-a a $y$ $\psi_\xi(y)$ is unambiguously defined on all of $I(\xi)$ We define $\psi$ $Y \to X$ by

$$\psi(y) = \psi_\xi(y) \quad \text{where } y \in Y(\xi) \quad \text{i e } \hat{y} = \xi$$

$\psi$ is finitary since for sufficiently large even $r$ $\psi(y)$ is determined on $I(\beta_r(\hat{y}))$, and hence at 0, by $y(J(\beta_r(\hat{y})))$ and $\pi_{\beta_r(\hat{y})}$ both of which are determined by the restriction of $y$ to the domain of $S_r(\hat{y})$, a finite portion of $y$ (However, see the remark after the end of this proof )

Finally for $B \subset X \times Y$ we obtain

$$\pi(B) = \nu\{y \in Y \quad (\psi(y), y) \in B\} \tag{v}$$

in the same way that we obtained (ii) (v) implies that $\psi$ is measure-preserving (ii) and (v) together imply that $\psi = \phi^{-1}$ for $E \supset X$

$$\begin{aligned}\mu\{\phi^{-1}\psi^{-1}E \cap E\} &= \mu\{x \in X \quad (x, \phi(x)) \in E \times \psi^{-1}E\} \\ &= \pi(E \times \psi^{-1}E) \\ &= \nu\{y \quad (\psi y, y) \in E \times \psi^{-1}E\} \\ &= \nu(\psi^{-1}(E)) = \mu(E)\end{aligned}$$

Since $\psi \circ \phi$ is measure-preserving this means $\mu(E \Delta \phi^{-1}\psi^{-1}E) = 0$, so $\psi\phi = id$ a e

□

*Remark.* Referring to the remark after the proof of Lemma 4 5 it is not hard to see that the $\pi_S$ are all causal Recall that

$$\pi_{\beta_r(\xi)} = \nu_{I_{-1}} * \pi_i * \pi_{i-1} * \qquad * \pi_j$$

where $\#(I_{-1}) = m_r$ and $i - j = M_r$, and $\pi_i, \quad, \pi_j$ are products of lower rank $\pi_{\mathscr{S}}$, and hence all causal The dual version of Lemma 4 5 then implies that the marginal of $\pi_{\beta_r(\xi)}$ on $A^{I(\beta_r(\xi))\cap(-\infty,0]} \times B^{J(\beta_r(\xi))\cap-(\infty\ 0]}$, which we denote $^*\pi_{\beta_r(\xi)}$, has the form

$$^*\pi_{\beta_r(\xi)} = \nu_{I_{-1}} * \pi_i * \pi_{i-1} * \qquad * {}^*\pi_t,$$

where $^*\pi_t$ denotes the marginal of $\pi_t$ on the past Thus Proposition 4 8 applies to $^*\pi_{\beta_r(\xi)}$ as well as $\pi_{\beta_r(\xi)}$ This leads to the observation that we can determine $\psi_\xi(y)(-\infty, 0]$ if we know $y(-\infty, 0)$ and $^*\pi_{\beta_r(\xi)}$ for all $r$ Since $^*\pi_{\beta_r(\xi)}$ is certainly determined by $\xi$ it follows that $\psi$ is marker-conditionally causal in the sense that once $\xi = \hat{y}$ is known then the past of $\psi(y)$ depends only on the past of $y$ However, one needs to know all of $\xi$ because $^*\pi_{\beta_r(\xi)}$ is not determined by the past of $\xi$ alone Indeed, since auxiliary and principal skeleta are determined by working from right to left, if we lose the right end of $\mathscr{S}_r(\xi)$ we will not even know what the domain of $^*\pi_{\beta_r(\xi)}$ is

## 6 *Proof of Theorem* 2

The purpose of this section is to sketch a proof of Theorem 2, the non-finitary version of Theorem 1 Technical details will for the most part be suppressed We assume familiarity with §§ 2 and 4

THEOREM 2 *There exısts a forgetful ısomorphısm* $\phi$ $X \to Y$ *such that* $\phi$ $X(\xi) \to Y(\xi)$

We contınue to assume that $p(1) = q(1)$ and we maıntaın all the notatıon ıntroduced ın §§ 2 and 4 Also extend that notatıon as follows If $I$ ıs finıte we ıdentıfy the partıtıon $P^I$ (or $Q^I$) on any space havıng $A^I$ as a factor wıth the finıte $\sigma$-algebra ıt generates If $I$ ıs ınfinıte $P^I$ denotes the $\sigma$-algebra generated by the projectıon $X \to A^I$, $A^I$ beıng gıven the usual product Borel structure We recall also one notatıonal eccentrıcıty from § 3 when $x$ ıs a sequence ındexed by $\mathbb{Z}$ and $I \subset \mathbb{Z}$, $x(I)$ denotes the restrıctıon of $x$ to $I$

If $(Z, \mathscr{B}, \lambda)$ ıs a probabılıty space, $\mathscr{F}$ ıs a sub-$\sigma$-algebra of $\mathscr{B}$ and $R$ ıs a finıte $\mathscr{B}$-measurable partıtıon of $Z$ we wrıte $R \subset \mathscr{F}$ $(\lambda)$ ıf each $r \in R$ agrees a e wıth an $r' \in \mathscr{F}$ For $\eta > 0$ we wrıte $R \overset{\varepsilon}{\subset} \mathscr{F}$ $(\lambda)$ ıf for each $r \in R$ there ıs an $r' \in \mathscr{F}$ such that $\lambda(r \,\Delta r') < \eta$ It ıs not hard to see that ıf $Q$ and $R$ are finıte partıtıons of $Z$ and

$$\lambda(\bigcup \{q \in Q \;\; \exists r \in R \text{ s t } \; q \subset r \; \lambda\text{-a e}\}) > 1 - \eta$$

then $R \overset{\eta}{\subset} \sigma(Q)$ $(\lambda)$ $(\sigma(Q))$ denotes the $\sigma$-algebra generated by $Q$)

To prove Theorem 2 we must construct a measurably varyıng famıly $\{\phi_\xi \;\; \xi \in \hat{X}\}$ of measurable and measure-preservıng mappıngs $\phi_\xi$ $X(\xi) \to Y(\xi)$ $\phi_\xi$ corresponds to a joınıng $\pi_\xi$ of $\mu_\xi$ and $\nu_\xi$ for $B \subset X(\xi) \times Y(\xi)$

$$\pi_\xi(B) = \mu_\xi\{x \in X(\xi) \;\; (x, \phi_\xi(x)) \in B\}$$

It ıs not hard to see that the famıly $\{\pi_\xi \;\; \xi \in \hat{X}\}$ wıll have to have the followıng propertıes for $\hat{\mu}$-a a $\xi$

(6 0) $\forall$ finıte cylınders $E \subset X \times Y$, the mappıng $\xi \mapsto \pi_\xi(E \cap (X(\xi) \times Y(\xi))$ ıs $\hat{\mu}$-measurable

(6 1) shıft ınvarıance $\pi_{\hat{\sigma}(\xi)} = (\sigma \times \tau)\pi_\xi$

(6 2) $\pi_\xi$ ıs a superposıtıon, ı e a joınıng of $\mu_\xi$ and $\nu_\xi$

(6 3) each $\pi_\xi$ ıs forgetful, that ıs $P^{-(\infty, j) \cap I(\xi)} \perp P^{[j, \infty) \cap I(\xi)} Q^{[j,\infty) \cap I(\xi)}$ $(\pi_\xi)$

(6 4) the margınal $\pi^*_\xi$ of $\pi_\xi$ on $A^{I(\xi)}$ depends only on $\xi[0, \infty)$

(6 3) ıs a consequence of the fact that $P^{[j,\infty) \cap I(\xi)} \supset Q^{[j,\infty) \cap I(\xi)}$ $(\pi_\xi)$, whıch follows from the forgetful nature of $\phi$ (6 4) ıs lıkewıse a consequence of the fact that $\phi$ ıs forgetful However, certaınly neıther (6 3) nor (6 4) ımplıes that $\{\pi_\xi\}$ arıses from a mappıng consıder, for example, $\pi_\xi = \mu_\xi \times \nu_\xi$

The facts that $\phi_\xi$ ıs a forgetful homomorphısm and ıs a e one-to-one are expressed by

(6 5) $$Q^{\{0\} \cap I(\xi)} \subset P^{[0, \infty) \cap I(\xi)} \quad (\pi_\xi)$$

and

(6 6) $$P^{\{0\} \cap I(\xi)} \subset Q^{I(\xi)} \quad (\pi_\xi)$$

(Of course these are vacuous when $0 \notin I(\xi)$) Conversely one checks that any famıly $\{\pi_\xi \;\; \xi \in \hat{X}\}$ satısfyıng (ı)-(vı) arıses from a $\phi$ as ın Theorem 2 In checkıng that $\phi$ ıs forgetful (6 4) ıs essentıal ıf we dıd not have ıt we mıght need the full marker sequence $\xi$ to determıne $\pi^*_\xi$, whıch ıs needed to determıne $\phi(x)(0)$ from $x[0, \infty)$

Accordıngly, we turn our attentıon from mappıngs to joınıngs satısfyıng (6 0)-(6 6) Let $M_\xi$ denote the space of Borel probabılıty measures on $X(\xi) \times Y(\xi)$. We denote by $M$ the space of $\hat{\mu}$-equivalence classes of functions $\pi\ \hat{X}\to\bigcup\{M_\xi\ \ \xi\in\hat{X}\}$, denoted $\xi\mapsto\pi_\xi$, such that $\pi_\xi\in M_\xi$ and (6 0)-(6 4) hold for $\hat{\mu}$-a a $\xi$ Note that $M$ is not empty let $\pi_\xi=\mu_\xi\times\nu_\xi$

For $\xi\in\hat{X}$ we define a complete metric $d_\xi$ on $M_\xi$ which induces the weak-* topology on $M_\xi$ as follows Let $\Gamma_i$ denote the set of all cylinder sets in $X\times Y$ depending only on coordinates in $[-i, i]$ For $B\subset X\times Y$ let $B_\xi=B\cap X(\xi)\times Y(\xi)$ and for $m^1, m^2\in M_\xi$ let

$$d_\xi(m^1, m^2)=\sum_{i=0}^{\infty}2^{-i}\sup_{C\in\Gamma_i}|m^1(C_\xi)-m^2(C_\xi)|$$

Now we define a complete metric $d$ on $M$ by

$$d(\pi^1,\pi^2)=\int_{\hat{X}}d_\xi(\pi^1_\xi,\pi^2_\xi)\,d\mu(\xi)$$

Thus $\pi^1$ and $\pi^2$ are close if $\pi^1_\xi$ and $\pi^2_\xi$ are weak-* close with high $\hat{\mu}$-probability

We introduce approximate versions of (6 5) and (6 6) by defining

$$\mathcal{U}_\varepsilon=\{\pi\in M\ \ \hat{\mu}\{\xi\ \ Q^{\{0\}\cap I(\xi)}\overset{\varepsilon}{\subset}P^{[0\ \infty)\cap I(\xi)}\quad(\pi_\xi)\}>1-\varepsilon\}$$

and

$$\mathcal{V}_\varepsilon=\{\pi\in M\ \ \hat{\mu}\{\pi\ \ P^{\{0\}\cap I(\xi)}\overset{\varepsilon}{\subset}Q^{I(\xi)}\quad(\pi_\xi)\}>1-\varepsilon\}$$

It is straightforward to check that $\mathcal{U}_\varepsilon$ and $\mathcal{V}_\varepsilon$ are open subsets of $M$ and that $\bigcap_{n\geq1}(\mathcal{U}_{1/n}\cap\mathcal{V}_{1/n})$ consists precisely of those $\pi\in M$ satisfying (6 5) and (6 6) Thus Theorem 2 is a consequence of the following proposition and the Baire category theorem

PROPOSITION 6 1 *The sets $\mathcal{U}_\varepsilon$ and $\mathcal{V}_\varepsilon$ are dense in $M$*

*Proof* We deal with $\mathcal{U}_\varepsilon$ first What we have to show is that for any $\pi\in M$ we can find $\tilde{\pi}\in\mathcal{U}_\varepsilon$ such that $d_\xi(\pi_\xi,\tilde{\pi}_\xi)$ is small for most $\xi$ We can ensure that $\tilde{\pi}\in\mathcal{U}_\varepsilon$ by ensuring that for most (that is, more than $1-\eta$ in $\hat{\mu}$-measure) $\xi$ we have

$$Q^{\{0\}\cap I(\xi)}\overset{\varepsilon}{\subset}P^{[0\ \infty)\cap I(\xi)}\quad(\tilde{\pi}_\xi)\tag{1}$$

(The point is that we will in fact be able to make $\eta$ small independent of $\varepsilon$)

Roughly speaking we will produce $\tilde{\pi}_\xi$ by combining the marginals of $\pi_\xi$ over large disjoint finite chunks of $\mathbb{Z}$ using the *-product, in a way which takes advantage of Proposition 4 7 to ensure that (1) holds for most $\xi$ We use the structure of $\xi$ to determine the chunks, ensuring shift-invariance of $\tilde{\pi}$ If the chunks are sufficiently large then for most $\xi$ a large interval $[-n, n]$ of $\mathbb{Z}$ will be contained in a single chunk so that $\tilde{\pi}_\xi$ will have exactly the same marginal (not just close) as $\pi_\xi$ over $[-n, n]$, whence $d_\xi(\pi_\xi,\tilde{\pi}_\xi)$ is small

To this end let $N_1<N_2$, $C$ and $m$ be integers to be specified later Suppose $\xi\in\hat{X}$ and $I=[i, j]\subset\mathbb{Z}$ is an interval such that

$$\xi(i-1)=0\quad\xi(i)=\xi(i+1)=\quad=\xi(j)=1\quad\xi(j+1)=0$$

$I$ will be called a 1-marker in $\xi$ if $N_1\leq j-i+1$ $I$ will be called a 2-marker if $N_2\leq j-i+1$ (so a 2-marker is also a 1-marker) Thus a a $\xi\in\hat{X}$ contain infinitely many 1-markers and infinitely many 2-markers By an $i$-skeleton ($i=1$ or 2) in $\xi$

we mean the set of all indices $t \in \mathbb{Z}$ such that $\xi(t) = 0$ which lie between two successive $i$-markers in $\xi$ (Note that this is formally different from the definition of a 1- or 2-skeleton in § 3 ) Thus for a a $\xi \in \hat{X}$ $I(\xi)$ decomposes disjointly into 2-skeleta each of which decomposes into 1-skeleta If $\gamma$ is a 1-skeleton in $\xi$, $C(\gamma)$ will denote its $C$ rightmost indices and $C(\gamma) = C(\gamma) - \{\max \gamma\}$

Now fix $\xi \in \hat{X}$ and suppose $\mathcal{S}$ is a 2-skeleton occurring in $\xi$ whose component 1-skeleta are $\gamma_t, \gamma_{t-1}$, , $\gamma_0$ listed in order from left to right We set

$$I(\mathcal{S}) = \mathcal{S}$$

and

$$J(\mathcal{S}) = C(\gamma_t) \cup C(\gamma_{t-1}) \cup \quad \cup C(\gamma_m)$$

(Thus $J(\mathcal{S}) = \varnothing$ if $t < m$) For $i \geq m$ let $\pi_i$ denote the marginal of $\pi_\xi$ on $A^{\gamma_i} \times \mathcal{B}^{C(\gamma_i)}$, for $i < m$ let $\pi_i$ denote $\mu_{\gamma_i}$ and let

$$\begin{aligned} \hat{\pi}_{\mathcal{S}} &= \pi_t * \pi_{t-1} * \quad * \pi_m * \pi_{m-1} * \quad * \pi_0 \\ &= \pi_t * \pi_{t-1} * \quad * \pi_m * \mu_{\gamma_{m-1} \cup \quad \cup \gamma_0} \end{aligned} \tag{ii}$$

Next we 'fill in the holes' of $\hat{\pi}_{\mathcal{S}}$ by letting $J' = \mathcal{S} - J(\mathcal{S})$ and defining

$$\tilde{\pi}_{\mathcal{S}} = \hat{\pi}_{\mathcal{S}} \times \nu_{J'}, \tag{iii}$$

a superposition on $A^{\mathcal{S}} \times \mathcal{B}^{\mathcal{S}}$ Finally if

$$, \mathcal{S}_{-1}, \mathcal{S}_0, \mathcal{S}_1, \mathcal{S}_2,$$

denote the 2-skeleta of $\xi$, listed in order of appearance, we set

$$\tilde{\pi}_\xi = \quad \tilde{\pi}_{\mathcal{S}_{-1}} \times \tilde{\pi}_{\mathcal{S}_0} \times \tilde{\pi}_{\mathcal{S}_1} \times \tilde{\pi}_{\mathcal{S}_2}$$

(It is not crucial that we take the product joining here In fact, roughly speaking, any 'canonical' joining of the $\pi'_{\mathcal{S}_i}$ which is a superposition, for example $\tilde{\pi}_{\mathcal{S}_{-1}} * \tilde{\pi}_{\mathcal{S}_0} * \tilde{\pi}_{\mathcal{S}_1}$ would do )

It is not difficult to check using Lemmas 4 2 and 4 4 that $\tilde{\pi} \in M$ We mention only that (6 4) is a consequence of (6 3), Lemma 4 5 and the fact that all our constructions are made working from right to left

Moreover, given $\eta > 0$ and $k \in \mathbb{Z}^+$, if $N_1$ is sufficiently large then, with $\hat{\mu}$-probability more than $1 - \eta$, 0 is enclosed in a 1-skeleton $\gamma = \gamma(\xi)$ of $\xi$, that is

$$0 \in [\min \gamma(\xi), \max \gamma(\xi)]$$

and $\gamma(\xi) \supset [-k, k+1] \cup I(\xi)$ If we next choose $C$ sufficiently large then with high $\hat{\mu}$-probability $C(\gamma(\xi)) = \gamma(\xi)$ so

$$C(\gamma(\xi)) \supset [-k, k] \cap I(\xi) \tag{iv}$$

Now assuming $\gamma(\xi)$ is defined let $\mathcal{S}(\xi)$ denote the 2-skeleton in $\xi$ containing $\gamma(\xi)$ and let $\mathcal{S}^*(\xi) = \mathcal{S}(\xi) \cap [0, \infty)$ Suppose that the 1-skeleta of $\mathcal{S}(\xi)$ are $\gamma_t, \gamma_{t-1}$, , $\gamma_0$ and $\gamma(\xi) = \gamma_{\bar{t}}$ so

$$\mathcal{S}^*(\xi) = \gamma_{\bar{t}}^* \cup \gamma_{\bar{t}-1} \cup \quad \cdot \quad \cup \gamma_0$$

where $\gamma_{\bar{t}}^* = \gamma_{\bar{t}} \cap [0, \infty)$ Let $J^*(\mathcal{S}(\xi)) = J(\mathcal{S}(\xi)) \cap [0, \infty)$ and let $\tilde{\pi}^*_{\mathcal{S}(\xi)}$ denote the marginal of $\tilde{\pi}_\xi$ on $A^{\mathcal{S}^*(\xi)} \times \mathcal{B}^{J^*(\mathcal{S}(\xi))}$, which is the same as the marginal of $\tilde{\pi}_{\mathcal{S}(\xi)}$ on this set, where $\tilde{\pi}_{\mathcal{S}(\xi)}$ is defined by (iii) with $\mathcal{S} = \mathcal{S}(\xi)$ Lemmas 4 4 and 4 5 together with the forgetfulness of $\pi_\xi$ imply that

$$\tilde{\pi}^*_{\mathcal{S}(\xi)} = \pi^*_{\bar{i}} * \pi_{\bar{i}-1} * \quad * \pi_m * \mu_{\gamma_{m-1}\cup \quad \cup\gamma_0},$$

where the $\pi_i$ are the measures appearing in (ii) and $\pi^*_{\bar{i}}$ denotes the marginal of $\pi_{\bar{i}}$ on $A^{\gamma^*_{\bar{i}}} \times B^{C(\gamma_{\bar{i}})\cap[0\ \infty)}$ Recalling that $\pi_i$ is a measure on $A^{\gamma_i} \times \mathcal{B}^{C(\gamma_i)}$, $\#C(\gamma_i) \leq C$ and $\#(\gamma_{m-1}\cup \quad \cup\gamma_0) \geq m$, Proposition 4 7 implies that if we choose $m$ sufficiently large then for all $\xi$ for which $\gamma(\xi)$ is defined we have

$$\mu_{\mathcal{S}^*(\xi)}(\bigcup\{p \in P^{\mathcal{S}^*(\xi)} \quad p \text{ is not split by } \tilde{\pi}_{\mathcal{S}(\xi)^*}\}) > 1-\varepsilon$$

so

$$Q^{J^*(\mathcal{S}(\xi))} \overset{\varepsilon}{\subset} P^{\mathcal{S}^*(\xi)} \quad (\tilde{\pi}^*_{\mathcal{S}(\xi)})$$

When the partitions are considered as partitions of $A^{I(\xi)} \times \mathcal{B}^{J(\xi)}$ this implies

$$Q^{J^*(\mathcal{S}(\xi))} \overset{\varepsilon}{\subset} P^{I(\xi)\cap[0,\infty)} \quad (\tilde{\pi}_\xi) \tag{v}$$

Finally, if $N_2$ is chosen sufficiently large then with high $\hat{\mu}$-probability $\gamma(\xi)$ is not among the rightmost $m$ 1-skeleta in $\mathcal{S}(\xi)$ so $C(\gamma(\xi)) \subset J(\mathcal{S}(\xi))$ Recalling the definition of $\tilde{\pi}_{\mathcal{S}(\xi)}$ this implies that with high probability $\tilde{\pi}_\xi$ and $\pi_\xi$ have the same marginal on $A^{\gamma(\xi)} \times \mathcal{B}^{C(\gamma(\xi))}$ In view of (iv) this implies that $d_\xi(\tilde{\pi}_\xi, \pi_\xi)$ is small Moreover (iv) together with $C(\gamma(\xi)) \subset J(\mathcal{S}(\xi))$ implies that $\{0\} \cap I(\xi) \subset J^*((\xi))$ so in view of (v) we obtain that with high $\hat{\mu}$-probability

$$Q^{\{0\}\cap I(\xi)} \overset{\varepsilon}{\subset} P^{I(\xi)\cap[0\ \infty)} \quad (\tilde{\pi}_\xi),$$

as desired This concludes the proof that $\mathcal{U}_\varepsilon$ is open

To show that $\mathcal{V}_\varepsilon$ is open we proceed as follows Fixing $\pi \in M$ we seek $\tilde{\pi} \in \mathcal{V}_\varepsilon$ which is close to $\pi$ Let $N_1 < N_2$, $C$, $m$ and $M$ be positive integers to be specified later Fix $\xi \in \hat{X}$ and define 1- and 2-skeleta in $\xi$ as before, as well as $\mathcal{S}(\xi)$ and $\gamma(\xi)$ For $\gamma$ a 1-skeleton $C(\gamma)$ and $C(\gamma)$ are also defined as before Suppose $\mathcal{S}$ is a 2-skeleton in $\xi$ with component 1-skeleta $\gamma_t, \gamma_{t-1}, \quad, \gamma_0$ For $0 \leq i \leq t$ let

$$i = q(M+m) + r(i) \quad 0 \leq r(i) < M+m$$

Call $\gamma_i$ principal if $0 \leq r(i) < M$ and auxiliary if $M \leq r(i) < M+m$ Define

$$I(\mathcal{S}) = \bigcup\{C(\gamma_i) \quad \gamma_i \text{ principal}\}$$

and

$$\pi_i = \begin{cases} \text{marginal of } \pi_\xi \text{ on } A^{C(\gamma_i)} \times \mathcal{B}^{\gamma_i}, \text{ if } \gamma_i \text{ principal} \\ \nu_{\gamma_i}, \text{ if } \gamma_i \text{ auxiliary} \end{cases}$$

Next let

$$\hat{\pi}_{\mathcal{S}} = \pi_t * \pi_{t-1} * \quad * \pi_0, \tag{vi}$$

a superposition on $A^{I(\mathcal{S})} \times \mathcal{B}^{\mathcal{S}}$, and fill in the holes of $\hat{\pi}_{\mathcal{S}}$ by defining

$$\tilde{\pi}_{\mathcal{S}} = \hat{\pi}_{\mathcal{S}} \times \mu_{\mathcal{S}-I(\mathcal{S})},$$

a superposition on $A^{\mathcal{S}} \times \mathcal{B}^{\mathcal{S}}$ Finally if

$$\mathcal{S}_{-1}, \mathcal{S}_0, \mathcal{S}_1, \mathcal{S}_2$$

are the component 2-skeleta of $\xi$ set

$$\tilde{\pi}_\xi = \quad \times \tilde{\pi}_{\mathcal{S}_{-1}} \times \tilde{\pi}_{\mathcal{S}_0} \times \tilde{\pi}_{\mathcal{S}_{-1}} \times$$

We can now argue, much as for $\mathcal{U}_\varepsilon$, that $\tilde{\pi} \in \mathcal{V}_\varepsilon$ and is close to $\pi$ $N_1$, $C$, $m$, $M$ and $N_2$ are chosen in that order The only significantly different feature here is how we ensure that for most $\xi$

$$P^{\{0\}\cap I(\xi)} \overset{\varepsilon}{\subset} Q^{I(\xi)} \quad (\tilde{\pi}_\xi) \tag{vii}$$

We begin, as with $\mathcal{U}_\varepsilon$, by choosing $N_1$ so large that for most $\xi$ $\gamma(\xi)$ is defined After $C$ and $m$ are chosen (we will have something to say about them later) if we choose $M$ and then $N_2$ sufficiently large then, for most $\xi$, $\gamma(\xi)$ lies in a string of $M$ consecutive principal 1-skeleta $\gamma_{i+M-1}, \gamma_{i+M-2}, \quad , \gamma_i$ of $\mathcal{S}(\xi)$ which is preceded by a string of $m$ auxiliary 1-skeleta $\gamma_{i+M+m-1}, \quad , \gamma_{i+M}$ For such a $\xi$ we set

$$\beta(\xi) = \gamma_j \cup \gamma_{j-1} \cup \quad \cup \gamma_{i+M} \cup \gamma_{i+M-1} \cup \quad \cup \gamma_i,$$

where $j = i + M + m - 1$ and

$$I(\beta(\xi)) = I(\mathcal{S}(\xi)) \cap \beta(\xi) = C(\gamma_{i+M-1}) \cup \quad \cup C(\gamma_i)$$

Now by associativity of $*$ the marginal of $\tilde{\pi}_\xi$ on $A^{I(\beta(\xi))} \times B^{\beta(\xi)}$, which we denote $\tilde{\pi}_{\beta(\xi)}$, has the form

$$\begin{aligned}\tilde{\pi}_{\beta(\xi)} &= (\pi_j * \pi_{j-1} * \quad * \pi_{i+M}) * \pi_{i+M-1} * \quad * \pi_i \\ &= \nu_{\gamma_j \cup \quad \cup \gamma_{i+M}} * \pi_{i+M-1} * \quad * \pi_i\end{aligned}$$

where the $\pi_i$ are the measures appearing in (vi) Thus Lemma 4 8 guarantees that if we chose $m$ sufficiently large then, regardless of how $M$ and $N_2$ were subsequently chosen,

$$P^{I(\beta(\xi))} \overset{\varepsilon}{\subset} Q^{\beta(\xi)} \quad (\tilde{\pi}_{\beta(\xi)})$$

so

$$P^{I(\beta(\xi))} \overset{\varepsilon}{\subset} Q^{I(\xi)} \quad (\tilde{\pi}_\xi) \tag{viii}$$

Now, as in the argument for density of $\mathcal{U}_\varepsilon$, the choice of $N_1$ and $C$ could have been made to ensure that for most $\xi$

$$C(\gamma(\xi)) \supset [-k, k] \cap I(\xi),$$

which ensures the closeness of $\pi_\xi$ and $\tilde{\pi}_\xi$ for most $\xi$ As we already have that $\gamma(\xi)$ is principal for most $\xi$ it also ensures that $\{0\} \cap I(\xi) \subset I(\beta(\xi))$ for most $\xi$ so (viii) implies that (vii) holds for most $\xi$ □

## REFERENCES


[B, R] R Burton & A Rothstein Isomorphism theorem in ergodic theory Preprint

[J] A del Junco Finitary codes between one-sided Bernoulli shifts *Ergod Th & Dynam Sys* **1** (1981), 285–301

[K, S] M Keane & M Smorodinsky Bernoulli shifts of the same entropy are finitarily isomorphic *Ann Math* **109** (1979), 397–406

[M] L Meshalkin A case of isomorphy of Bernoulli schemes *Dokl Akad Nauk SSSR* **128** (1959), 41–44

[O] D Ornstein *Ergodic Theory, Randomness and Dynamical Systems* New Haven and London, Yale University Press (1974)

[O, W] D Ornstein & B Weiss Unilateral codings of Bernoulli systems *Israel J Math* **21** (1975), 159–166

[P] J Propp PhD dissertation, University of California, Berkeley (1987)

[S] Ja G Sinai A weak isomorphism of transformations with an invariant measure *Dokl Akad Nauk SSSR* **147** (1962), 797–800

[W] P Walters Some results on the classification of non-invertible measure-preserving transformations *Lecture Notes in Mathematics* **318**, pp 266–276 Springer Verlag, Berlin (1973)