STUDIA MATHEMATICA 188 (1) (2008)

Epsilon-independence between two processes

by

Tomasz Downarowicz and Paulina Grzegorek (Wrocław)

Abstract. We study the notion of ε-independence of a process on finitely (or countably) many states and that of ε-independence between two processes defined on the same measure preserving transformation. For that we use the language of entropy. First we demonstrate that if a process is ε-independent then its ε-independence from another process can be verified using a simplified condition. The main direction of our study is to find natural examples of ε-independence. In case of ε-independence of one process, we find an example among processes generated on the induced (first return time) transformation defined on a typical long cylinder set of any given process of positive entropy. To obtain examples of pairs of ε-independent processes we have to make an additional assumption on the master process. Then again, we find such pairs generated on the induced transformation as above. This is the most elaborate part of the paper. While the question whether our assumption is necessary remains open, we indicate a large class of processes where our assumption is satisfied.

1. Introduction. Let $(X,\Sigma,\mu,T,\mathcal{P})$ be a process generated on a standard probability measure preserving invertible transformation (X,Σ,μ,T) by a finite measurable partition $\mathcal{P}$ of X. Such a process is called *independent* if the discrete random variables $\pi_n : X \to \mathcal{P}$ defined by the rule $\pi_n(x) = p \Leftrightarrow T^n(x) \in p$ $(n \in \mathbb{Z})$ are independent (i.e., every finite collection of such variables is independent). A synonym for an independent process is a *Bernoulli process*; however, some authors use the latter name for any process measure-theoretically isomorphic to an independent one.

A process $(X,\Sigma,\mu,T,\mathcal{P})$ is ε-independent if it satisfies an approximate independence condition (the precise definition will be provided later). Such processes play an important role in the Ornstein theory (see e.g. [Sh] for an exposition on variants of this notion). The key (and rather obvious, once the definition is stated) observation is that a process is independent if and only if it is ε-independent for every $\varepsilon > 0$.

2000 *Mathematics Subject Classification*: 37A35, 37A05, 60G10.

Key words and phrases: ergodic process, positive entropy, ε-independence, induced process, return time.

Now consider two finite partitions of X, say $\mathcal{P}$ and $\mathcal{Q}$. The meaning of the statement: "the processes $(X, \Sigma, \mu, T, \mathcal{P})$ and $(X, \Sigma, \mu, T, \mathcal{Q})$ are mutually independent" is clear: any finite collection of random variables of the form π_n defined by the first process should be jointly independent of any finite collection of analogous random variables defined by the latter process. It seems interesting to introduce a meaningful notion of ε-independence between two processes generated by two partitions on the same measure preserving transformation. In a recent paper [D-L], a prototype of such notion has been introduced and it proved very useful in the study of the so-called "return time asymptotics". In the same paper the authors raise the question whether a more general ε-independence holds for some specific pairs of processes occurring naturally in every process of positive entropy. The details of this question will be discussed later in this note.

This paper addresses ε-independence of a single process and between two processes. We replicate from [D-L] an example of a class of ε-independent processes naturally occurring in any process of positive entropy. Such a process exists on the induced system (with the action of the first return map) on "nearly every" sufficiently long cylinder B; it is simply the process generated by the partition $\mathcal{P}$ restricted to B. Also, we propose the notions of "limit ε-independence" and "ε-independence" between two processes. We observe that two processes are mutually indpendent if and only if they are mutually ε-independent for every $\varepsilon > 0$. Further, we prove that if one of the processes is ε-independent then the notions of limit ε-independence and ε-independence coincide. Finally, in our main result, we prove that under some additional assumptions on the process $(X, \Sigma, \mu, T, \mathcal{P})$, the following two processes defined on the induced system on a typical long cylinder B are mutually ε-independent: the ε-independent process generated by the partition $\mathcal{P}$ restricted to B (the same as mentioned above), and the $\mathbb{N}$-valued process of return times to B. These are exactly the processes appearing in the question formulated in [D-L]. Due to the additional assumption, we do not claim to have solved the problem posed in [D-L], nevertheless, we indicate a large class of systems with positive entropy which satisfy our additional condition.

2. Notation and preliminaries. Let (X, Σ, μ, T) be an automorphism of a standard probability measure space. Let $\mathcal{P}$ be a finite or countable measurable partition of the space X. We will use the following notation: if $\mathbb{A} \subset \mathbb{Z}$ then

$$\mathcal{P}^{\mathbb{A}} = \mathcal{P}^{T,\mathbb{A}} = \bigvee_{i \in \mathbb{A}} T^{-i}(\mathcal{P}). \tag{2.1}$$

We will skip the first superscript only when it indicates the transformation T; we will keep it with respect to any other transformation. We will abbreviate $\mathcal{P}^n = \mathcal{P}^{[0,n-1]}$ and $\mathcal{P}^{-n} = \mathcal{P}^{[-n,-1]}$. The sub-$\sigma$-field $\mathcal{P}^{\pm}$ generated by the union of the partitions $T^n(\mathcal{P})$ over all $n \in \mathbb{Z}$ is invariant, i.e., $T(\mathcal{P}^{\pm}) = \mathcal{P}^{\pm}$, and the corresponding factor system is referred to as the *process generated by* $\mathcal{P}$. The sub-σ-field generated by the union of the partitions $T^n(\mathcal{P})$ $(n \geq 1)$ will be denoted by $\mathcal{P}^-$ and called the *past of the process*. Clearly, it is *subinvariant*, i.e., $T(\mathcal{P}^-) \subset \mathcal{P}^-$. In case it is necessary to point out the transformation, the symbols $\mathcal{P}^{T,n}, \mathcal{P}^{T,\pm}$, etc. will appear.

The process generated by $\mathcal{P}$ is often represented in terms of symbolic dynamics, as follows. Let $X_{\mathcal{P}} = \mathcal{P}^{\times\mathbb{Z}}$ be the set of all doubly infinite sequences (x_n) with values in $\mathcal{P}$. Let $\Sigma_{\mathcal{P}}$ denote the product σ-field in $X_{\mathcal{P}}$. Let σ denote the left shift transformation on $X_{\mathcal{P}}$, $\sigma(x_n) = (x_{n+1})$. Then there is a natural (measurable) map $\pi : X \to X_{\mathcal{P}}$, $\pi(x) = (x_n)$ defined by the rule $x_n = p \in \mathcal{P} \Leftrightarrow T^n(x) \in p$. It is customary to call $\pi(x)$ the $\mathcal{P}$*-name* of x. We obviously have $\pi \circ T = \sigma \circ \pi$. Finally, we obtain a shift-invariant measure $\mu_{\mathcal{P}}$ on $\Sigma_{\mathcal{P}}$ from μ "via preimage": for $B \in \Sigma_{\mathcal{P}}$ we set $\mu_{\mathcal{P}}(B) = \mu(\pi^{-1}(B))$. The process generated by $\mathcal{P}$ is isomorphic to $(X_{\mathcal{P}}, \Sigma_{\mathcal{P}}, \mu_{\mathcal{P}}, \sigma)$. In this setup, the elements of $\mathcal{P}^n$ coincide with the (lifted by π) cylinders over the blocks B of length n over $\mathcal{P}$. Here $\mathcal{P}$ is treated as the *alphabet* (just a set of labels or *symbols*), the block B has the form $B = B[0, n-1] = [b_0, b_1, \ldots, b_{n-1}] \in \mathcal{P}^{\times n}$, the cylinder over B is the set $C_B = \{(x_n) \in X_{\mathcal{P}} : x[0, n-1] = B\}$ and the lifted cylinder is $\pi^{-1}(C_B)$, or equivalently, $\{x \in X : T^i(x) \in b_i,\ i = 0, 1, \ldots, n-1\}$. Similarly, the elements of $\mathcal{P}^{-n}$ correspond to the lifted cylinders of the form $\{x \in X : T^i(x) \in b_{i+n},\ i = -n, -n+1, \ldots, -1\}$. To simplify the notation, we will denote the lifted cylinders by the same letters as the underlying blocks and write $B \in \mathcal{P}^n$ or $B \in \mathcal{P}^{-n}$. In this manner both $\mathcal{P}^n$ and $\mathcal{P}^{-n}$ formally denote $\mathcal{P}^{\times n}$, the nth cartesian power of $\mathcal{P}$, but we keep track of the positioning of the blocks on the time axis while defining the cylinders.

Given a countable (or finite) partition α of X, its *entropy* (or *static entropy*) is defined by

$$H(\alpha) = H(\mu, \alpha) = -\sum_{A \in \alpha} \mu(A) \log(\mu(A)), \tag{2.2}$$

with the convention $0 \log 0 = 0$. Notice that on the set of partitions of cardinality bounded by some M, this function depends continuously on the values $\mu(A)$, moreover, it is strictly concave. If the measure μ is fixed, we will write $H(\alpha)$, while $H(\mu, \alpha)$ is used when the indication of the measure involved is necessary.

Suppose that β is another countable partition. Then the *conditional entropy of* α *given* β is defined as

$$H(\alpha \,|\, \beta) = H(\mu, \alpha \,|\, \beta) = \sum_{B \in \beta} \mu(B) H(\mu_B, \alpha), \tag{2.3}$$

where μ_B is the conditional measure on B: $\mu_B(A) = \mu(A \cap B)/\mu(B)$ (if $\mu(B) = 0$, μ_B is inessential for the formula, so it can be defined as any probability measure). Since for any A,

$$\mu(A) = \sum_{B \in \beta} \mu(B)\mu_B(A), \tag{2.4}$$

strict concavity of H implies that $H(\alpha \,|\, \beta) \le H(\alpha)$, and equality holds if and only if $\mu_B(A) = \mu(A)$ for every B and A, i.e., when the partitions α and β are stochastically independent.

If $H(\beta) < \infty$, the following alternative formula for the conditional entropy holds:

$$H(\alpha \,|\, \beta) = H(\alpha \vee \beta) - H(\beta). \tag{2.5}$$

Notice, by the way, that always $H(\alpha) \le H(\alpha \vee \beta) \le H(\alpha) + H(\beta)$.

If $\mathcal{F} \subset \Sigma$ is a sub-σ-field, then the conditional entropy of α given $\mathcal{F}$ is defined as

$$H(\alpha \,|\, \mathcal{F}) = H(\mu, \alpha \,|\, \mathcal{F}) = \inf H(\alpha \,|\, \beta), \tag{2.6}$$

where β ranges over all $\mathcal{F}$-measurable countable partitions. It is not hard to prove that the infimum is achieved along any sequence of finite partitions β_n such that β_{n+1} refines β_n for each n, and $\mathcal{F}$ is the σ-field generated by the union of the partitions β_n.

We now recall basic facts about the dynamical entropy of a partition. If (X, Σ, μ, T) is a probability measure preserving transformation and $\mathcal{P}$ is a finite partition of X then we let

$$H_n(\mathcal{P}) = H_n(\mu, \mathcal{P}) = \frac{1}{n} H(\mathcal{P}^n), \tag{2.7}$$

and

$$h(\mathcal{P}) = h(\mu, \mathcal{P}) = h(\mu, T, \mathcal{P}) = \lim_{n\to\infty} H_n(\mathcal{P}) = H(\mathcal{P} \,|\, \mathcal{P}^-). \tag{2.8}$$

We will tend to use the short version with respect to the "master" transformation T and measure μ. Other transformations or measures will be shown in the notation. The above limit exists by the well-known monotonicity of the sequence $H_n(\mathcal{P})$, and the last equality is also well-known and easy to prove.

If $\mathcal{Q}$ is another (possibly countable) partition, then we define

$$H_n(\mathcal{P} \,|\, \mathcal{Q}) = H_n(\mu, \mathcal{P} \,|\, \mathcal{Q}) = \frac{1}{n} H(\mathcal{P}^n \,|\, \mathcal{Q}^n), \tag{2.9}$$

and

$$h(\mathcal{P} \,|\, \mathcal{Q}) = h(\mu, \mathcal{P} \,|\, \mathcal{Q}) = h(\mu, T, \mathcal{P} \,|\, \mathcal{Q}^{T,\pm}) = \lim_{n\to\infty} H_n(\mathcal{P} \,|\, \mathcal{Q}). \tag{2.10}$$

This limit exists by (relatively easy to verify) subadditivity of the sequence $H(\mathcal{P}^n \,|\, \mathcal{Q}^n)$. The sequence $H_n(\mathcal{P} \,|\, \mathcal{Q})$ is in fact nonincreasing, but this fact is almost unknown and has a rather complicated proof, so we will not invoke it. We will be needing the following: if $\mathcal{P}$ is finite then

$$h(\mathcal{P}\,|\,\mathcal{Q}) = \lim_{n\to\infty}\frac{1}{n}H(\mathcal{P}^n\,|\,\mathcal{Q}^{\pm}) = H(\mathcal{P}\,|\,\mathcal{P}^{-}\vee\mathcal{Q}^{\pm}) \tag{2.11}$$

(see e.g. [D-S] for a proof). If, in addition, $H(\mathcal{Q})<\infty$, then

$$h(\mathcal{P}\,|\,\mathcal{Q}) = h(\mathcal{P}\vee\mathcal{Q}) - h(\mathcal{Q}). \tag{2.12}$$

3. Independence and ε-independence. We will soon make repeated use of an obvious fact concerning weighted averages (or, more generally, integrals with respect to probability measures). For easy reference we isolate this fact and call it the *rectangle rule* (we skip the easy proof).

FACT 3.1. *Let ξ be a probability measure on a space Ω and let $f\le g$ be two real (measurable) functions on Ω. If $\int f\,d\xi > \int g\,d\xi - \gamma\delta$, where $\gamma>0$ and $\delta>0$, then $\xi\{\omega\in\Omega : f(\omega)\le g(\omega)-\gamma\}<\delta$.*

Two countable (including finite) partitions α and β are *ε-independent* (we write $\alpha \perp\!\!\!\perp^{\varepsilon} \beta$) if

$$\sum_{A\in\alpha,\,B\in\beta} |\mu(A\cap B)-\mu(A)\cdot\mu(B)| < \varepsilon. \tag{3.1}$$

Note that the above holds for all $\varepsilon>0$ if and only if the partitions are stochastically independent.

The connection between entropy and independence is captured by the following fact.

FACT 3.2. *For every $M\in\mathbb{N}$ and $\varepsilon>0$ there is $\delta>0$ such that for any finite partition α with at most M elements and any countable partition β, the following implications hold*:

$$\alpha \perp\!\!\!\perp^{\delta} \beta \ \Rightarrow\ H(\alpha\,|\,\beta) > H(\alpha)-\varepsilon, \tag{3.2}$$

and

$$H(\alpha\,|\,\beta) > H(\alpha)-\delta \ \Rightarrow\ \alpha \perp\!\!\!\perp^{\varepsilon} \beta. \tag{3.3}$$

In particular, independence between α and β is equivalent to the equality

$$H(\alpha\,|\,\beta) = H(\alpha) \tag{3.4}$$

(as already observed earlier).

Proof. Assume δ-independence. Reformulation of (3.1) yields

$$\sum_{B\in\beta}\mu(B)\sum_{A\in\alpha}|\mu_B(A)-\mu(A)| < \delta \tag{3.5}$$

By the rectangle rule (for $f(B) \equiv 0$, $g(B) = \sum_{A\in\alpha} |\mu_B(A) - \mu(A)|$), for a collection of B's of joint measure at least $1 - \sqrt{\delta}$ we have

$$\sum_{A\in\alpha} |\mu_B(A) - \mu(A)| < \sqrt{\delta}. \tag{3.6}$$

By continuity of H on partitions of cardinality M, and by compactness of the simplex of M-dimensional probability vectors, for such B's we have $|H(\mu_B, \alpha) - H(\mu, \alpha)| < \varepsilon/2$, if δ is a priori (independently of α, depending only on M) chosen small enough. Finally,

$$\begin{aligned} H(\alpha \,|\, \beta) = \sum_{B\in\beta} \mu(B) H(\mu_B, \alpha) &\ge (1 - \sqrt{\delta})(H(\mu, \alpha) - \varepsilon/2) \\ &> H(\alpha) - \varepsilon, \end{aligned} \tag{3.7}$$

again, for a well chosen δ.

Conversely, by strict concavity of H on the compact set of M-dimensional probability vectors there is a uniform choice of δ such that $\#\alpha < M$ yields

$$\sum_{B\in\beta} \mu(B) \sum_{A\in\alpha} |\mu_B(A) - \mu(A)| \ge \varepsilon \;\Rightarrow\; H(\alpha \,|\, \beta) \le H(\alpha) - \delta. \tag{3.8}$$ ■

Fact 3.2 allows one to use the condition $H(\alpha \,|\, \beta) > H(\alpha) - \varepsilon$ as an alternative notion of ε-independence for a finite partition. We will then say that α *is ε-entropy independent of* β. If both partitions have finite entropy, the condition can be written as $H(\alpha \vee \beta) > H(\alpha) + H(\beta) - \varepsilon$, hence this is a symmetric relation.

We say that a finite partition α is *ε-independent of a σ-field $\mathcal{F}$* if it is ε-independent of any countable $\mathcal{F}$-measurable partition β. Analogously, ε-entropy independence of $\mathcal{F}$ is defined by $H(\alpha \,|\, \mathcal{F}) > H(\alpha) - \varepsilon$.

The process $(X, \Sigma, \mu, T, \mathcal{P})$ generated by a finite partition $\mathcal{P}$ is called an *independent process* if the partitions $\{T^n(\mathcal{P}) : n \in \mathbb{Z}\}$ are jointly independent. By invariance, an equivalent condition is that $\mathcal{P}$ is independent of the past $\mathcal{P}^-$, which is equivalent to $h(\mathcal{P}) = H(\mathcal{P})$.

Definition 3.3. The process is called *ε-independent* (resp. *ε-entropy independent*) if $\mathcal{P}$ is ε-independent (resp. ε-entropy independent) of $\mathcal{P}^-$.

ε-entropy independence of a process can be written as

$$h(\mathcal{P}) > H(\mathcal{P}) - \varepsilon, \tag{3.9}$$

or $H_n(\mathcal{P}) > H(\mathcal{P}) - \varepsilon$ for all n.

Remark 3.4. Clearly, if $H(\mathcal{P})$ is smaller than ε then the generated process is (trivially) ε-entropy independent. It is thus natural to require, for nontriviality of the notion, that ε is much smaller than $H(\mathcal{P})$.

REMARK 3.5. If a process is ε-entropy independent for every $\varepsilon > 0$ then it is an independent process.

We now turn to the case of two partitions $\mathcal{P}$ and $\mathcal{Q}$.

DEFINITION 3.6. Assume that $\mathcal{P}$ is finite, while we let $\mathcal{Q}$ be countable. We will say that the process $(X, \Sigma, \mu, T, \mathcal{P})$ generated by a partition $\mathcal{P}$ is *ε-entropy limit-independent* of the process generated by $\mathcal{Q}$ if

$$h(\mathcal{P} \mid \mathcal{Q}) > h(\mathcal{P}) - \varepsilon. \tag{3.10}$$

An equivalent condition is that eventually (for large n),

$$H_n(\mathcal{P} \mid \mathcal{Q}) > H_n(\mathcal{P}) - \varepsilon' \tag{3.11}$$

for some $\varepsilon' > \varepsilon$. If, in addition, $\mathcal{Q}$ has finite entropy, then the last inequality can be equivalently written as

$$H_n(\mathcal{P} \vee \mathcal{Q}) > H_n(\mathcal{P}) + H_n(\mathcal{Q}) - \varepsilon', \tag{3.12}$$

which proves that in that case the relation is symmetric. Notice that ε-entropy limit-independence between two processes for every $\varepsilon > 0$ does not imply full stochastic independence.

REMARK 3.7. Clearly, a process of entropy smaller than ε is (trivially) ε-entropy limit-independent from any process (even from itself).

This is why we introduce a stronger notion.

DEFINITION 3.8. The process generated by $\mathcal{P}$ is *ε-entropy independent* of the process generated by $\mathcal{Q}$ if

$$H_n(\mathcal{P} \mid \mathcal{Q}) > H_n(\mathcal{P}) - \varepsilon \tag{3.13}$$

for every n.

This notion is also symmetric among finite entropy partitions. Now, ε-entropy independence for every $\varepsilon > 0$ does imply stochastic independence between the processes. As before, for the notion to be nontrivial it is required that ε is smaller than both $H(\mathcal{P})$ and $H(\mathcal{Q})$.

The notions of ε-entropy independence and ε-entropy limit-independence coincide (via change of the parameter) if one of the processes is itself an ε-entropy independent process.

FACT 3.9. *Suppose $(X, \Sigma, \mu, T, \mathcal{P})$ is an ε-entropy independent process and that it is ε-entropy limit-independent of the process generated by another partition $\mathcal{Q}$. Then the former process is 2ε-entropy independent of the latter.*

Proof. $H_n(\mathcal{P} \mid \mathcal{Q}) \geq h(\mathcal{P} \mid \mathcal{Q}) > h(\mathcal{P}) - \varepsilon > H(\mathcal{P}) - 2\varepsilon \geq H_n(\mathcal{P}) - 2\varepsilon$. ∎

4. Induced processes. Both the notion of an ε-independent process and that of a pair of mutually ε-independent processes are meaningful only if there exist natural classes of examples for each of them. Moreover, we would like to have (separate) examples for any arbitrarily small ε, and in which ε-independence does not result from small entropy or full stochastic independence. We will provide such examples appearing naturally "inside" any process of positive entropy. For this we need some preliminaries on so-called induced maps.

As usual, let (X, Σ, μ, T) denote a probability measure preserving transformation. For a set $B \in \Sigma$ of positive measure we define $R_B : B \to \mathbb{N}$ as the *first return time* to B:

$$R_B(x) = \min\{n > 0 : T^n(x) \in B\}. \tag{4.1}$$

By the Poincaré theorem, R_B is finite μ_B-almost everywhere. The *induced map* $T_B : B \to B$ is defined by

$$T_B(x) = T^{R_B(x)}(x). \tag{4.2}$$

It is easy to see that T_B preserves the measure μ_B. The Abramov theorem asserts that the measure-theoretic entropies: $h_\mu(T)$ (of the measure preserving transformation T with respect to μ) and $h_{\mu_B}(T_B)$ (of T_B with respect to μ_B) are bound by the simple relation

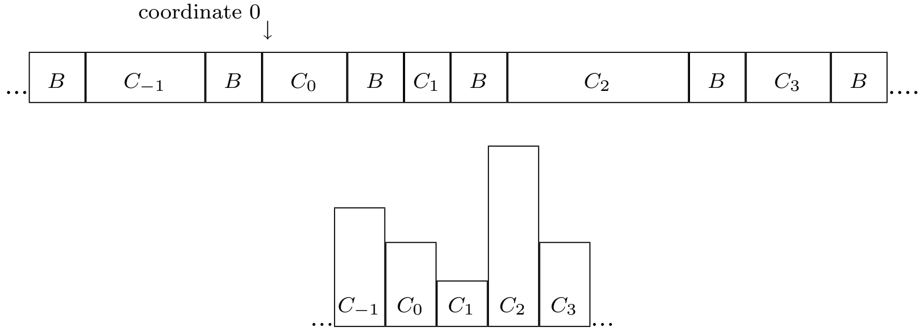
$$h_{\mu_B}(T_B) = h_\mu(T)/\mu(B). \tag{4.3}$$

Recall that, by the Sinai theorem, the measure-theoretic entropy $h_\mu(T)$ of a measure preserving transformation equals the dynamical entropy $h(\mu, T, \mathcal{P})$ of a process generated by a partition $\mathcal{P}$ as soon as $\mathcal{P}$ is a *generator* for T, i.e., when $(X, \Sigma, \mu, T, \mathcal{P})$ is isomorphic to (X, Σ, μ, T).

Suppose that $\mathcal{P}$ is a generator and that B is a cylinder in $\mathcal{P}^{-n}$ (i.e., a block occurring at positions $[-n, -1]$). Two points $x, y \in B$ are equal if and only if their $\mathcal{P}$-names (under the transformation T) are equal. Thus, in order to generate for the transformation T_B, a partition of B must distinguish (under T_B) the full $\mathcal{P}$-names. The simplest such partition is

$$\mathcal{Q}_B = \mathcal{P}^{[0, R_B - 1]}. \tag{4.4}$$

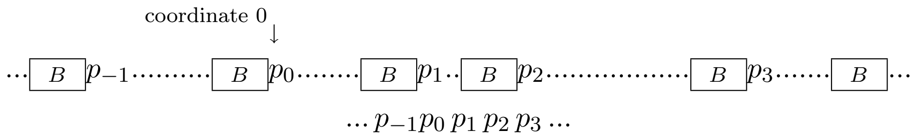
Notice that the "exponent" is variable (depends on a point x), which has to be properly understood. First, we take the partition $\mathcal{R}_B$ of B depending on the return time: $\mathcal{R}_B = \{R_n : n \geq 1\}, R_n = \{x \in B : R_B(x) = n\}$, then we refine the partition $\mathcal{R}_B$ intersecting every set R_n with the cylinders of length n. The partition $\mathcal{Q}_B$ is usually infinite (still countable), and its "symbols" correspond to the blocks appearing in between consecutive repetitions of the block B in the $\mathcal{P}$-names of the points of B. The figure below illustrates the passage between a symbolic representation of $x \in B$ in the process $(X, \Sigma, \mu, T, \mathcal{P})$ and in the induced process $(B, \Sigma_B, \mu_B, T_B, \mathcal{Q}_B)$.



Because, as we said, $\mathcal{Q}_B$ is a generator for T_B, the Abramov formula yields

$$h(\mu_B, T_B, \mathcal{Q}_B) = h(\mu, T, \mathcal{P})/\mu(B). \tag{4.5}$$

Notice that $\mathcal{P}$ is usually not a generator for the induced system. The generated process $(B, \Sigma_B, \mu_B, T_B, \mathcal{P})$ "reads" only the single symbols that follow directly to the right of the occurrences of B, as illustrated in the figure below.



There is no simple direct formula for the corresponding entropy $h(\mu_B, T_B, \mathcal{P})$. Clearly, it is usually much smaller than $h(\mu, T, \mathcal{P})/\mu(B)$, comparable with $h(\mu, T, \mathcal{P})$. As the next theorem says, for n large enough, a typical such process is ε-independent.

Let us introduce one more convention that will help us avoid frequent repetitions of a lengthy expression. We will say that a property $\Phi(B)$ holds *with μ-tolerance ε* for n-cylinders $B \in \mathcal{P}^{-n}$ if the measure of the union of all cylinders $B \in \mathcal{P}^{-n}$ for which Φ fails to hold does not exceed ε.

Theorem 4.1. *Let $(X, \Sigma, \mu, T, \mathcal{P})$ be a process generated by a finite partition $\mathcal{P}$ and let $\varepsilon > 0$ be given. Then, for n sufficiently large, with μ-tolerance ε the n-cylinders $B \in \mathcal{P}^{-n}$ have the property that the process $(X, \Sigma, \mu_B, T_B, \mathcal{P})$ is ε-entropy independent.*

Proof. Because $H(\mathcal{P}\,|\,\mathcal{P}^{-n}) \searrow H(\mathcal{P}\,|\,\mathcal{P}^{-})$, for large n we have $H(\mathcal{P}\,|\,\mathcal{P}^{-n}) - \varepsilon^2 < H(\mathcal{P}\,|\,\mathcal{P}^{-})$. Now

$$\begin{aligned}\sum_{B\in\mathcal{P}^{-n}} \mu(B) H(\mu_B, \mathcal{P}) - \varepsilon^2 &= H(\mathcal{P}\,|\,\mathcal{P}^{-n}) - \varepsilon^2 < H(\mathcal{P}\,|\,\mathcal{P}^{-}) \\ &= H(\mathcal{P}\,|\,\mathcal{P}^{-n} \vee \mathcal{P}^{-}) = \sum_{B\in\mathcal{P}^{-n}} \mu(B) H(\mu_B, \mathcal{P}\,|\,\mathcal{P}^{-}).\end{aligned} \tag{4.6}$$

Clearly, for every B, the term $H(\mu_B,\mathcal{P})$ dominates $H(\mu_B,\mathcal{P}\,|\,\mathcal{P}^-)$. However, (4.6) shows that the weighted average of the first terms exceeds the weighted average of the latter terms by no more than ε^2. Thus, by the rectangle rule,

$$H(\mu_B,\mathcal{P}) < H(\mu_B,\mathcal{P}\,|\,\mathcal{P}^-)+\varepsilon, \tag{4.7}$$

except for sets B of joint measure at most ε. We let X_n be the union of the cylinders B satisfying the above. Because the past of the induced process is contained (as a σ-field) in the full past $\mathcal{P}^-$, we have

$$H(\mu_B,\mathcal{P}\,|\,\mathcal{P}^-) \le H(\mu_B,\mathcal{P}\,|\,\mathcal{P}^{T_B,-}) = h(\mu_B,T_B,\mathcal{P}) \tag{4.8}$$

for every B. Thus for every B contained in X_n, $H(\mu_B,\mathcal{P}) < h(\mu_B,T_B,\mathcal{P})+\varepsilon$, i.e., the process generated by $\mathcal{P}$ for the induced map is ε-entropy independent, as claimed. ∎

In [D-L] the following question has been posed:

Question 1. Is it true that for the majority of sufficiently long cylinders $B\in\mathcal{P}^{-n}$, the (ε-entropy independent) process $(X,\Sigma,\mu_B,T_B,\mathcal{P})$ is also ε-entropy limit-independent (hence, by Fact 3.9, 2ε-entropy independent) of the process of return times $(X,\Sigma,\mu_B,T_B,\mathcal{R}_B)$?

Remark 4.2. The question is whether

$$h(\mu_B,T_B,\mathcal{P}\,|\,\mathcal{R}_B) > h(\mu_B,T_B,\mathcal{P})-\varepsilon, \tag{4.9}$$

i.e., whether

$$H(\mu_B,\mathcal{P}\,|\,\mathcal{P}^{T_B,-}\vee\mathcal{R}_B^{T_B,\pm}) > H(\mu_B,\mathcal{P}\,|\,\mathcal{P}^{T_B,-})-\varepsilon. \tag{4.10}$$

In [D-L] the following partial result is proved: Given $K\in\mathbb{N}$ and $\varepsilon>0$, for the majority of sufficiently long cylinders B,

$$H(\mu_B,\mathcal{P}\,|\,\mathcal{P}^{T_B,-}\vee\mathcal{R}_B^{T_B,(-\infty,K]}) > H(\mu_B,\mathcal{P}\,|\,\mathcal{P}^{T_B,-})-\varepsilon. \tag{4.11}$$

Remark 4.3. Notice that in a process of positive entropy neither $h(\mu_B,T_B,\mathcal{P})$ nor $h(\mu_B,T_B,\mathcal{R}_B)$ decrease to zero on average (over all blocks of given length) as the length grows. For this reason, Question 1 cannot be answered trivially using Remark 3.7. In fact both average entropies are bounded below by $h(\mu,T,\mathcal{P})$.

Indeed, it follows from calculations (4.6) and (4.8) that

$$\sum_{B\in\mathcal{P}^{-n}}\mu(B)h(\mu_B,T_B,\mathcal{P})\ge h(\mu,T,\mathcal{P}).$$

To see the inequality for the process of return times, consider the partition $\{B,B^c\}$. Notice that the process generated by this partition and then induced on B is precisely $(X,\mu_B,T_B,\mathcal{R}_B)$, hence $h(\mu,T,\{B,B^c\})=$

$\mu(B)h(\mu_B, T_B, \mathcal{R}_B)$. Summing over all $B \in \mathcal{P}^{-n}$ we get

$$\begin{aligned}\sum_{B\in\mathcal{P}^{-n}} \mu(B)h(\mu_B, T_B, \mathcal{R}_B) &= \sum_{B\in\mathcal{P}^{-n}} h(\mu, T, \{B, B^c\}) \\ &\ge h\Big(\mu, T, \bigvee_{B\in\mathcal{P}^{-n}} \{B, B^c\}\Big) = h(\mu, T, \mathcal{P}^{-n}) \\ &= h(\mu, T, \mathcal{P}).\end{aligned}$$

5. The main result. In this section we will provide a positive answer to Question 1 under an additional assumption on the process. We believe that our technique will allow in the future to skip the assumption or prove that it is always satisfied. A simple example of a process that satisfies it without being independent or even weakly mixing is given at the end of the paper.

THEOREM 5.1. *Let $(X, \Sigma, \mu, T, \mathcal{P})$ be an ergodic process with positive entropy $h = H(\mathcal{P} \,|\, \mathcal{P}^-)$ such that*

$$\forall_{\delta>0}\ \exists_{\gamma\in(0,1/4)}\ \exists_{C>1}\ \exists_{M(\delta)}\ \forall_{n\ge M(\delta)}$$

there exists a partition $\mathcal{F}_n \cup \{Z_n\} \prec \mathcal{P}^{-n}$ satisfying

$$\mu(Z_n) < \delta, \tag{5.1}$$

$$\forall_{F,F'\in\mathcal{F}_n} \quad \mu(F)/\mu(F') < C, \tag{5.2}$$

and, if F_n denotes the element of $\mathcal{F}_n$ of maximal measure, then

$$\forall_{n\ge M(\delta)} \quad 1 + \delta < \mu(F_n)/\mu(F_{n+1}) < C, \tag{5.3}$$

$$\forall_{F\in\mathcal{F}_n} \quad \mu_F\{x : R_F(x) < 2\gamma/\mu(F_n) + n\} < \delta. \tag{5.4}$$

Then for every $\varepsilon > 0$ there exists S such that for every $s \ge S$ with μ-tolerance ε for $B \in \mathcal{P}^{-s}$ the processes $(X, \Sigma, \mu_B, T_B, \mathcal{P})$ and $(X, \Sigma, \mu_B, T_B, \mathcal{R}_B)$ are ε-entropy limit-independent, i.e.,

$$h(\mu_B, T_B, \mathcal{P} \,|\, \mathcal{R}_B) > h(\mu_B, T_B, \mathcal{P}) - \varepsilon. \tag{5.5}$$

Proof. Fix $\varepsilon \in (0, 1)$. Choose

$$\delta < \frac{\varepsilon^2}{6h} \quad \text{and} \quad \zeta < \frac{\gamma\varepsilon^2}{24C^2}, \tag{5.6}$$

where γ and C depend on δ as in the assumption of the theorem. By a standard argument involving the Shannon–McMillan–Breiman theorem (see e.g. [P]) there exists m such that

$$H(\mu_E, \mathcal{P}^m) \le m(h + \zeta) \tag{5.7}$$

for any subset $E \subset X$ of measure larger than $1/2$. We can also assume that

$$\frac{\log 2}{m} < \frac{\zeta}{2}. \tag{5.8}$$

Choose $N > \max\{M(\delta), m\}$ such that for each $n \geq N$ the following conditions are satisfied:

$$h \leq H(\mathcal{P} \mid \mathcal{P}^{-n}) < h + \zeta, \tag{5.9}$$

$$m\mu(F_n) \log(\#\mathcal{P}) < \zeta/4, \tag{5.10}$$

$$\frac{\gamma}{C\mu(F_n)} > n. \tag{5.11}$$

Now we define $S = \lfloor \gamma/\mu(F_N) + N \rfloor$. Then for every $s > S$ we find $n_s > N$ such that

$$\frac{\gamma}{\mu(F_{n_s-1})} + n_s - 1 \leq s < \frac{\gamma}{\mu(F_{n_s})} + n_s. \tag{5.12}$$

We choose γ_s smaller than γ so that $s - n_s = \gamma_s/\mu(F_{n_s}) =: q_s$. Note that (5.11) and the assumption (5.3) imply

$$q_s = s - n_s \geq \frac{\gamma}{\mu(F_{n_s-1})} - 1 > \frac{\gamma}{C\mu(F_{n_s})} - 1 > n_s - 1, \tag{5.13}$$

hence

$$q_s \geq n_s. \tag{5.14}$$

Now we estimate the ratio between γ and γ_s. Notice that γ_s is the smallest when the following equality holds (recall conditions (5.3) and (5.11)):

$$\frac{\gamma_s}{\mu(F_{n_s})} + n_s = \frac{\gamma}{\mu(F_{n_s-1})} + n_s - 1. \tag{5.15}$$

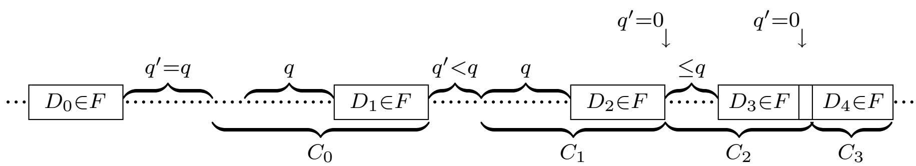
Then

$$\gamma_s = \gamma \frac{\mu(F_{n_s})}{\mu(F_{n_s-1})} - \mu(F_{n_s}) \geq \frac{\gamma}{C} - \frac{\gamma}{2C} = \frac{\gamma}{2C}. \tag{5.16}$$

Main calculations. At this point we fix some $s > S$, $n = n_s$, $q = q_s$ and $F \in \mathcal{F}_n$. The entropy of the complete induced process on F satisfies

$$\begin{aligned} \frac{h}{\mu(F)} &= h(\mu_F, T_F, \mathcal{Q}_F) \\ &= h(\mu_F, T_F, \mathcal{P}^{[q', R_F - 1]}) + h(\mu_F, T_F, \mathcal{P}^{q'} \mid \mathcal{P}^{[q', R_F - 1]}) \\ &=: W_1 + W_2 \end{aligned} \tag{5.17}$$

(recall (4.4) and (4.5); they apply to unions of cylinders of the same length, as well), where $q' = \max\{\min\{q, R_F - s\}, 0\}$. Note that $q' = q'(x)$ becomes a variable. The figure below shows q' in four different cases depending on whether $R_F \geq 2q + n$, $R_F \in (q + n, 2q + n)$, $R_F \in [n, q + n]$, $R_F < n$.



By (5.14) and (5.4),

$$\mu_F(\{q' < q\}) < \delta. \tag{5.18}$$

We now estimate W_1 from above, which combined with (5.17) will give us a lower estimate of W_2.

The elements of $\mathcal{P}^{[q',R_F-1]}$ are cylinders C_k of various lengths appearing in the $\mathcal{P}$-names of points $x \in X$ starting q' positions to the right from every occurrence of a block $D_k \in F$ and ending at the right end of the next block $D_{k+1} \in F$ (see figure above).

Let $E = \{T^i x : x \in F,\ q'(x) < i \le R_F(x)\}$ (a point belongs here if the coordinate -1 is covered by a block C_k). The process $(E, \Sigma_E, \mu_E, T_E, T(\mathcal{P}))$ "reads one by one" the symbols in the blocks C_k, "jumping" over the gaps between them. Note also that $F \subset E$ and the full induced process on F obtained from $(E, \Sigma_E, \mu_E, T_E, T(\mathcal{P}))$ is $(F, \Sigma_F, \mu_F, T_F, \mathcal{P}^{[q',R_F-1]})$. Hence

$$W_1 = h(\mu_F, T_F, \mathcal{P}^{[q',R_F-1]}) = \frac{\mu(E)}{\mu(F)}\, h(\mu_E, T_E, T(\mathcal{P})). \tag{5.19}$$

Note that by (5.18),

$$\mu(E) = 1 - \mu(F) \int_F q'\, d\mu_F \le 1 - \mu(F) q(1-\delta), \tag{5.20}$$

so

$$W_1 \le \left(\frac{1}{\mu(F)} - q(1-\delta)\right) h(\mu_E, T_E, T(\mathcal{P})). \tag{5.21}$$

Further, $h(\mu_E, T_E, T(\mathcal{P})) \le H_m(\mu_E, T_E, T(\mathcal{P})) = m^{-1} H(\mu_E, (T(\mathcal{P}))^{T_E,m})$. Observe that $(T(\mathcal{P}))^{T_E,m}$ differs from $(T(\mathcal{P}))^{T,m}$ only on the set $\bigcup_{i=0}^{m-1} T^{-i}(F)$ $=: V$, and its measure is at most $m\mu(F)$, which, by (5.10), is (much) smaller than $1/4$. Continuing,

$$\begin{aligned} H(\mu_E, (T(\mathcal{P}))^{T_E,m}) &\le H(\mu_E, (T(\mathcal{P}))^{T_E,m} \mid \{V, V^c\}) + H(\mu_E, \{V, V^c\}) \\ &\le H(\mu_{E\cap V^c}, (T(\mathcal{P}))^{T_E,m}) + \mu_E(V) H(\mu_{E\cap V}, (T(\mathcal{P}))^{T_E,m}) + \log 2 \\ &\le H(\mu_{E\cap V^c}, (T(\mathcal{P}))^{T,m}) + m\mu_E(F) \log(\#\mathcal{P}^m) + \log 2. \end{aligned} \tag{5.22}$$

Notice that since $\mu(E) \ge 1 - q\mu(F) \ge 1 - \gamma > 3/4$, the measure $\mu(E \cap V^c)$ is greater than $1/2$. The first term on the right hand side of (5.22) equals $H(\mu_{T^{-1}(E\cap V^c)}, \mathcal{P}^m)$, which by (5.7) does not exceed $m(h+\zeta)$. The second term is at most $m\zeta/2$ (use (5.10) and the fact that $\mu_E(F) \le 2\mu(F)$). After dividing by m and applying (5.8), we obtain

$$h(\mu_E, T_E, T(\mathcal{P})) \le h + 2\zeta. \tag{5.23}$$

Using (5.21) we obtain

$$W_1 \le \left(\frac{1}{\mu(F)} - q(1-\delta)\right)(h + 2\zeta). \tag{5.24}$$

Hence, by elementary computations and (5.2), (5.6), (5.16),

$$W_2 \geq qh(1-\delta) - \frac{2\zeta}{\mu(F)} \geq qh(1-\delta) - \frac{2C\zeta}{\mu(F_n)} \geq q\Big(h - \frac{\varepsilon^2}{3}\Big). \tag{5.25}$$

We will now transform W_2 to see how it is related to the left hand side of the assertion (5.5). We have

$$\begin{aligned} W_2 = h(\mu_F, T_F, \mathcal{P}^{q'} \,|\, \mathcal{P}^{[q',R_F-1]}) &\leq h(\mu_F, T_F, \mathcal{P}^{q} \,|\, \mathcal{P}^{[q',R_F-1]}) \\ &= \sum_{i=0}^{q-1} h(\mu_F, T_F, T^{-i}\mathcal{P} \,|\, \mathcal{P}^{[q',R_F+i-1]}). \end{aligned}$$

Fix $0 \leq i < q$. The partition $T_F\mathcal{P}^{[q',R_F+i-1]}$ will be denoted by $\mathcal{G}$. Now,

$$\begin{aligned} h(\mu_F, T_F, T^{-i}\mathcal{P} \,|\, \mathcal{P}^{[q',R_F+i-1]}) &= h(\mu_F, T_F, T^{-i}\mathcal{P} \,|\, T_F^{-1}(\mathcal{G})) \\ &= h(\mu_F, T_F, T^{-i}\mathcal{P} \,|\, \mathcal{G}) \\ &= H(\mu_F, T^{-i}\mathcal{P} \,|\, (T^{-i}\mathcal{P})^{T_F,-} \vee \mathcal{G}^{T_F,-} \vee \mathcal{G} \vee \mathcal{G}^{T_F,+}) \\ &\leq H(\mu_F, T^{-i}\mathcal{P} \,|\, (T^{-i}\mathcal{P})^{T_F,-} \vee \mathcal{G}^{T_F,-} \vee \mathcal{P}^{[-s+i,i-1]} \vee \mathcal{G}^{T_F,+}) \\ &= \sum_{B\in\mathcal{P}^{[-s+i,i-1]}} \mu_F(B) H(\mu_B, T^{-i}\mathcal{P} \,|\, (T^{-i}\mathcal{P})^{T_F,-} \vee \mathcal{G}^{T_F,-} \vee \mathcal{G}^{T_F,+}). \end{aligned}$$

To justify the above inequality consider two cases. If $q' > 0$ then $R_F \geq q' + q + n = q' + s$ and hence $T_F^{-1}\mathcal{G} = \mathcal{P}^{[q',R_F+i-1]} \succ \mathcal{P}^{[R_F-s+i,R_F+i-1]} = T_F^{-1}\mathcal{P}^{[-s+i,i-1]}$. If $q' = 0$ then maybe $\mathcal{G} \not\succ \mathcal{P}^{[-s+i,i-1]}$, still the missing coordinates are contained in $\mathcal{G}^{T_F,-}$ and thus $\mathcal{G}^{T_F,-} \vee \mathcal{G} \succ \mathcal{G}^{T_F,-} \vee \mathcal{P}^{[-s+i,i-1]}$.

It is easy to see that the σ-algebra of all return times to B (i.e., $\mathcal{R}_B^{T_B,\pm}$) is coarser than $\mathcal{G}^{T_F,-} \vee \mathcal{G}^{T_F,+}$ and that $(T^{-i}\mathcal{P})^{T_B,-}$ is coarser than $(T^{-i}\mathcal{P})^{T_F,-} \vee \mathcal{G}^{T_F,-}$. Thus we have

$$\begin{aligned} &\sum_{B\in\mathcal{P}^{[-s+i,i-1]}} \mu_F(B) H(\mu_B, T^{-i}\mathcal{P} \,|\, (T^{-i}\mathcal{P})^{T_F,-} \vee \mathcal{G}^{T_F,-} \vee \mathcal{G}^{T_F,+}) \\ &\qquad \leq \sum_{B\in\mathcal{P}^{[-s+i,i-1]}} \mu_F(B) H(\mu_B, T^{-i}\mathcal{P} \,|\, (T^{-i}\mathcal{P})^{T_B,-} \vee \mathcal{R}_B^{T_B,\pm}) \\ &\qquad = \sum_{B\in\mathcal{P}^{[-s+i,i-1]}} \mu_F(B) h(\mu_B, T_B, T^{-i}\mathcal{P} \,|\, \mathcal{R}_B). \end{aligned}$$

We have obtained

$$W_2 \leq \sum_{i=0}^{q-1} \sum_{B\in\mathcal{P}^{[-s+i,i-1]}\cap F} \mu_F(B) h(\mu_B, T_B, T^{-i}\mathcal{P} \,|\, \mathcal{R}_B). \tag{5.26}$$

Combining (5.25) and (5.26) and averaging over all $F \in \mathcal{F}_n$, we get

$$\Big(\sum_{F\in\mathcal{F}_n} \mu(F)\Big) q(h-\varepsilon^2/3) \le \sum_{F\in\mathcal{F}_n} \mu(F) \sum_{i=0}^{q-1} \sum_{B\in\mathcal{P}^{[-s+i,i-1]}} \mu_F(B) h(\mu_B, T_B, T^{-i}\mathcal{P} \,|\, \mathcal{R}_B).$$

We can move the sumation over i to the left. Then we can add to the right hand side the contribution to the entropy from the (missing) sets $B \subset Z_n$. The sum is now over all blocks $B \in \mathcal{P}^{[-s+i,i-1]}$ and it no longer depends on i. Thus, we can write

$$\sum_{i=0}^{q-1} \sum_{F\in\mathcal{F}_n} \mu(F) \sum_{B\in\mathcal{P}^{[-s+i,i-1]}} \mu_F(B) h(\mu_B, T_B, T^{-i}\mathcal{P} \,|\, \mathcal{R}_B) \le q \sum_{B\in\mathcal{P}^{-s}} \mu(B) h(\mu_B, T_B, \mathcal{P} \,|\, \mathcal{R}_B)$$

Now we use the fact that $\mu(Z_n) < \delta$ to obtain

$$\begin{aligned} q(h-2\varepsilon^2/3) \le (1-\delta) q(h-\varepsilon^2/3) &\le \Big(\sum_{F\in\mathcal{F}_n} \mu(F)\Big) q(h-\varepsilon^2/3) \\ &\le q \sum_{B\in\mathcal{P}^{-s}} \mu(B) h(\mu_B, T_B, \mathcal{P} \,|\, \mathcal{R}_B) \\ &\le q \sum_{B\in\mathcal{P}^{-s}} \mu(B) h(\mu_B, T_B, \mathcal{P}) \\ &\le q \sum_{B\in\mathcal{P}^{-s}} \mu(B) H(\mu_B, \mathcal{P}) \\ &= qH(\mu, \mathcal{P} \,|\, \mathcal{P}^{-s}) \\ &\le q(h+\zeta) \qquad (5.27) \\ &\le q(h+\varepsilon^2/3) \end{aligned}$$

using (5.9) for the inequality (5.27) and (5.6) near the end of the computations. Thus, the difference between

$$\sum_{B\in\mathcal{P}^{-s}} \mu(B) h(\mu_B, T_B, \mathcal{P} \,|\, \mathcal{R}_B) \quad \text{and} \quad \sum_{B\in\mathcal{P}^{-s}} \mu(B) h(\mu_B, T_B, \mathcal{P})$$

is not larger than ε^2. Since $h(\mu_B, T_B, \mathcal{P} \,|\, \mathcal{R}_B) \le h(\mu_B, T_B, \mathcal{P})$ for each $B \in \mathcal{P}^{-s}$, we can apply the rectangle rule to conclude that with μ-tolerance ε for $B \in \mathcal{P}^{-s}$,

$$0 \le h(\mu_B, T_B, \mathcal{P}) - h(\mu_B, T_B, \mathcal{P} \,|\, \mathcal{R}_B) \le \varepsilon. \ \blacksquare$$

6. A class of examples. In this section we indicate a class of processes which satisfy the assumptions (and hence the assertion) of Theorem 5.1. We remark that, by a theorem of Sinai ([S]), any process of positive entropy is an extension of an independent process. Below we require that the extension $X_{\mathcal{P}}$ is via a finite code. Notice that any process $(X, \Sigma, \mu, T, \mathcal{P})$ is isomorphic to one which is an extension via a block code, even via an amalgamation, i.e., a code with radius $r = 0$ (see below for the meaning of r) of an independent process, but has a rather specific generator. For example, the partition $\mathcal{P} \vee \mathcal{Q}$, where $\mathcal{Q}$ is the independent generator of the Bernoulli factor whose existence is granted by the theorem of Sinai, generates a process isomorphic to the original one, while it maps to the independent process generated by $\mathcal{Q}$ by the amalgamation $P \cap Q \mapsto Q$ ($P \in \mathcal{P}$, $Q \in \mathcal{Q}$).

THEOREM 6.1. *Let (X, Σ, μ, T) be an invertible measure preserving transformation and let $\mathcal{P}$ and $\mathcal{Q}$ be two measurable partitions of X. Suppose $\mathcal{Q}$ generates an independent process, while $\mathcal{P}$ is such that there exists a finite code from $X_{\mathcal{P}}$ to $X_{\mathcal{Q}}$, i.e., there exists an $r \in \mathbb{N}$ such that $\mathcal{P}^{[-r,r]}$ refines $\mathcal{Q}$. Then the process $(X, \Sigma, \mu, T, \mathcal{P})$ satisfies the assumptions* (5.1)–(5.4) *of Theorem* 5.1.

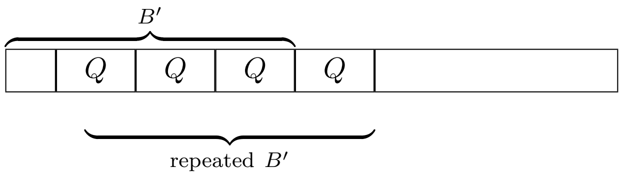
Proof. First notice that if we have cylinders of length n grouped into sets $F \in \mathcal{F}_n$ satisfying (5.1)–(5.4) in the independent process $(X, \Sigma, \mu, T, \mathcal{Q})$ then the preimages of the sets F via the finite code form a correct grouping of cylinders of length $n+2r$ in $(X, \Sigma, \mu, T, \mathcal{P})$. The conditions (5.1)–(5.3) are satisfied since the factor map preserves the measure, and (5.4) holds since the sets $F \in \mathcal{F}_n$ occur in $(X, \Sigma, \mu, T, \mathcal{Q})$ at exactly the same times as their preimages in $(X, \Sigma, \mu, T, \mathcal{P})$. (In order for the condition (5.4) to have correct parameters for the preimage cylinders, we must replace "$+\,n$" by "$+\,n + 2r$" in this condition.) Therefore it suffices to prove that the grouping is possible in any independent process.

Let $(X, \Sigma, \mu, T, \mathcal{Q})$ be an independent process of entropy h. Pick $\delta > 0$ so small that $e^{\frac{3}{4}(h-\delta/4)} > 1 + 2\delta$ and choose $\gamma < \delta/4$. We let $C = 2e^h$ (not depending on δ). We choose $M(\delta)$ so large that for every $n > M(\delta)$ the Shannon–McMillan–Breiman theorem is satisfied for $[n/2]$ with $\delta/4$, the expressions (6.1), (6.2) below are smaller than $\delta/4$, (6.4) is smaller than $\delta/2$, the right hand side of (6.5) is smaller than $\delta/2$, and (6.3) holds.

If B is a block of length n, we will denote by B' its "left half", i.e., the block $B[0, [n/2] - 1]$.

We begin by creating the set Z_n. First, we include in Z_n all cylinders B of length n which do not obey the S-M-B theorem or such that B' does not obey the S-M-B theorem. The joint measure of such cylinders is smaller than $2(\delta/4)$.

Next, we also put into Z_n the cylinders B of length n such that B' is repeated again within B. We need to estimate the joint measure of the cylinders satisfying this condition. If B is such a block then let $k \le [n/2]$ be the smallest integer such that $B[k, k+[n/2]-1] = B'$. Let $Q = B[[n/2], k+[n/2]-1]$. This block of length k completely determines B' (B' consists of periodic repetitions of Q except perhaps the leftmost repetition which may be truncated, see the diagram below).



Let $B'(Q)$ denote the block B' of length $[n/2]$ determined by $Q \in \mathcal{Q}^k$ in the above manner. The joint measure of all blocks B with the property that B' is repeated again, can be estimated from above by the following sum:

$$\sum_{Q\in\mathcal{Q}} \mu(B'(Q))\mu(Q) + \sum_{Q\in\mathcal{Q}^2} \mu(B'(Q))\mu(Q) + \cdots + 2\sum_{Q\in\mathcal{Q}^{[n/2]}} \mu(Q)\mu(Q)$$

$$\le ([n/2]+1)e^{-[n/2](h-\delta/4)} \tag{6.1}$$

(in the last sum $B'(Q) = Q$). Above, the measures of the cylinders $B'(Q)$ are estimated by $e^{-[n/2](h-\delta/4)}$, because the cylinders B for which B' does not satisfy the S-M-B-theorem have already been included in Z_n in the preceding step. The last sum is doubled to include the case when $k = [n/2]+1$. By the choice of n the measure of the set Z_n (so far) does not exceed $\frac{3}{4}\delta$.

The remaining cylinders will be grouped to form the sets $F \in \mathcal{F}_n$ (more precisely, some cylinders will still be added to Z_n). First we classify cylinders B into groups $\mathcal{G}(B')$ by their initial “half-block” B'. From each group $\mathcal{G}(B')$ we will create many sets $F \in \mathcal{F}_n$, as follows: We choose one by one different cylinders $B \in \mathcal{G}(B')$ (say in a random order) and we add their measures. We stop when the joint measure is for the first time not smaller than $e^{-\frac{3}{4}n(h-\delta/4)}$. The union of the chosen cylinders B is our first set $F \in \mathcal{F}$. Note that every cylinder B has measure at most $e^{-n(h-\delta/4)}$, much smaller than the measure of F, while the union of the whole group $\mathcal{G}(B')$ has measure at least $e^{-[n/2](h+\delta/4)}$, much larger than F. Thus, after F is created, in $\mathcal{G}(B')$ there are still many “unused” cylinders B left and we can similarly create the next set $F \in \mathcal{F}_n$, again adding cylinders from $\mathcal{G}(B')$ until the desired measure is reached. We repeat the construction until the joint measure of the unused cylinders B in $\mathcal{G}(B')$ is smaller than $e^{-\frac{3}{4}n(h-\delta/4)}$. Finally, we add these “left-over cylinders” to the set Z_n. We apply the same procedure with respect to every group $\mathcal{G}(B')$.

As a result, the measure of the set Z_n is increased by at most

$$e^{-\frac{3}{4}n(h-\delta/4)}e^{[n/2](h+\delta/4)} \tag{6.2}$$

since there are (according to the S-M-B theorem applied to B') at most $e^{[n/2](h+\delta/4)}$ groups $\mathcal{G}(B')$. We have arranged the number (6.2) to be smaller than $\delta/4$. Jointly, the measure of Z_n is smaller than δ.

Notice that the measure of any set $F \in \mathcal{F}_n$ is less than $e^{-\frac{3}{4}n(h-\delta/4)} + e^{-n(h-\delta/4)}$ and thus condition (5.2) is satisfied with the constant $C = 2$, and the more so with $C = 2e^h$. The right hand side inequality in condition (5.3) is satisfied by a straightforward verification and the left hand side inequality follows from the computation below:

$$\begin{aligned} \frac{\mu(F_n)}{\mu(F_{n+1})} &\geq \frac{e^{-\frac{3}{4}n(h-\delta/4)}}{e^{-\frac{3}{4}(n+1)(h-\delta/4)} + e^{-(n+1)(h-\delta/4)}} \\ &= \frac{e^{\frac{3}{4}(h-\delta/4)}}{1+e^{-\frac{n+1}{4}(h-\delta/4)}} > 1+\delta, \end{aligned} \tag{6.3}$$

where the last inequality holds by the choice of δ and n.

We now check the condition (5.4). Consider a set $F \in \mathcal{F}$. All points x in F have the same block B' at positions $[0, [n/2]-1]$. First we estimate the conditional measure μ_F of the event $\{R_F(x) < n\}$. This event happens only when B' is repeated in x before position n. But by the construction of F, B' does not occur in x before time $[n/2]$. Its occurrence at any position between $[n/2]$ and n is independent of its occurrence at coordinate 0, so we can write

$$\begin{aligned} \mu_F\{R_F(x) < n\} &\leq \frac{1}{\mu(F)}\left(\left[\frac{n}{2}\right]+1\right)e^{-2[n/2](h-\delta/4)} \\ &\leq \left(\left[\frac{n}{2}\right]+1\right)e^{-\frac{n}{4}(h-\delta/4)}. \end{aligned} \tag{6.4}$$

Finally, we estimate $\mu_F\{n \leq R_F < 2\gamma/\mu(F_n)+n+2r\}$. (Recall that $2r$ is added to the left bound, so that the condition (5.4) is properly carried over to preimages via the code of radius r.) By independence, this conditional measure does not exceed $2\gamma/\mu(F_n)+2r$ times the measure of F, i.e.,

$$2\gamma + 2r\mu(F) < 2\gamma + 4re^{-\frac{3}{4}n(h-\delta/4)}. \tag{6.5}$$

Now, $\mu_F\{R_F < 2\gamma/\mu(F_n)+n+2r\}$ is estimated by the sum of the right hand sides of (6.4) and (6.5), which we have arranged to be less than δ. The proof is complete. ■

References

[D-L] T. Downarowicz and Y. Lacroix, *The law of series*, submitted.

[D-S] T. Downarowicz and J. Serafin, *Fiber entropy and variational principles in compact non-metrizable spaces*, Fund. Math. 172 (2002), 217–242.

[P] K. Petersen, *Ergodic Theory*, Cambridge Univ. Press, 1983.

[Sh] P. Shields, *The Theory of Bernoulli Shifts*, Chicago Lectures in Math., Univ. of Chicago Press, Chicago, 1973.

[S] Ya. G. Sinai, *Weak isomorphism of transformations with an invariant measure*, Dokl. Akad. Nauk SSSR 147 (1962), 797–800 (in Russian); English transl.: Soviet Math. Dokl. 3 (1963), 1725–1729.

Institute of Mathematics and Computer Science
Wrocław University of Technology
Wybrzeże Wyspiańskiego 27
50-370 Wrocław, Poland
E-mail: downar@pwr.wroc.pl
paulina.grzegorek@pwr.wroc.pl

Received November 13, 2007
Revised version April 17, 2008 (6244)