Brudno's theorem for $\mathbb{Z}^d$ (or $\mathbb{Z}_+^d$) subshifts

Toru Fuda *and Miho Tonozaki †
Department of Mathematics, Hokkaido University
Sapporo 060-0810, Japan

Abstract

We generalize Brudno's theorem of 1-dimensional shift dynamical system to $\mathbb{Z}^d$ (or $\mathbb{Z}_+^d$) subshifts. That is to say, in $\mathbb{Z}^d$ (or $\mathbb{Z}_+^d$) subshift, the Kolmogorov-Sinai entropy is equivalent to the Kolmogorov complexity density almost everywhere for an ergodic shift-invariant measure.

Keywords. Brudno's theorem, Kolmogorov-Sinai entropy, Kolmogorov complexity, Shannon-McMillan-Breiman theorem, Subshifts, $\mathbb{Z}^d$-action, Universally typical sets

1 Introduction

In a topological dynamical system, A. A. Brudno defined a complexity of the trajectory of a point in the space by using the notion of Kolmogorov complexity, and showed the relationship between this quantity and the Kolmogorov-Sinai entropy [2]. As a preliminary step, Brudno considered the 1-dimensional shift dynamical system and showed that, for an ergodic shift-invariant measure, the Kolmogorov complexity density is equal to the Kolmogorov-Sinai entropy almost everywhere [2, Theorem 1.1].

A partial approach to generalize this theorem to a d-dimensional case is found in [10]. S. G. Simpson showed that, in $\mathbb{Z}^d$ (or $\mathbb{Z}_+^d$) subshifts, there exists a point such that its Kolmogorov complexity density is coincident with the topological entropy [10]. Examining Simpson's proof, we see that what he showed substantively is that the Kolmogorov complexity density is equal to the Kolmogorov-Sinai entropy almost everywhere only for a measure of maximal entropy.

The purpose of this paper is to generalize the Brudno's theorem of the $\mathbb{Z}_+^1$-action shift dynamical system to $\mathbb{Z}^d$ (or $\mathbb{Z}_+^d$) subshifts. The main theorem is the following:

*Corresponding author. E-mail: t-fuda@math.sci.hokudai.ac.jp
†Present affiliation: NEC Corporation, 5-7-2, Shiba, Minato, Tokyo, Japan.

Theorem 3.1 *If $\mu \in EM(S, \varsigma)$, then*

$$\mathcal{K}(\omega) = h_\varsigma(\mu), \quad \mu\text{-}a.e.\omega \in S.$$

Here S denotes $\mathbb{Z}^d$ (or $\mathbb{Z}_+^d$) subshift, ς denotes the shift action on S, $EM(S, \varsigma)$ denotes the set of all ergodic shift-invariant measures on the topological dynamical system (S, ς), $\mathcal{K}(\omega)$ denotes the Kolmogorov complexity density of ω, and $h_\varsigma(\mu)$ denotes the Kolmogorov-Sinai entropy of the measure preserving dynamical system $(S, \mathfrak{B}(S), \mu, \varsigma)$. We give the rigorous definition of these terms in Section 2.

In Section 2, we introduce some basic mathematical notions in ergodic theory, Kolmogorov complexity and shift dynamical systems. We used [4, 7, 9, 11] as main references for this section. Using these basic notions, we define the Kolmogorov complexity density of each point of $\Sigma^{\mathbb{Z}^d}$ (or $\Sigma^{\mathbb{Z}_+^d}$) naturally. In Section 3, we prove the main theorem and give some examples. The proof essentially uses Shannon-McMillan-Breiman therem and universally typical sets.

2 Some Mathematical Preliminaries

We first give quick reviews for some mathematical results related to the main theorem. We will not give proofs of theorems, see e.g. [4, 8]. We write $\mathbb{N} = \{1, 2, \cdots\}$, $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$, $\mathbb{Z}_+ = \{0, 1, 2, \cdots\}$. For an arbitrary fixed $d \in \mathbb{N}$, we set $G := \mathbb{Z}^d$ or $G := \mathbb{Z}_+^d$. For all $n \in \mathbb{N}$, let $\Lambda_n := \{g = (g_i)_{i=1}^d \in G : \forall i \in \{1, \cdots, d\}, |g_i| < n\}$. Then we have

$$|\Lambda_n| = \begin{cases} (2n-1)^d & (G = \mathbb{Z}^d), \\ n^d & (G = \mathbb{Z}_+^d), \end{cases}$$

where we denote by $|A|$ the cardinality of a set A.

2.1 Ergodic theory

Let $(X, \mathfrak{B}, \mu, \mathcal{T})$ be a *measure preserving dynamical system* (m.p.d.s.), namely, $(X, \mathfrak{B}, \mu)$ be a probability space and $\mathcal{T} = (T^g)_{g \in G}$ be a measurable μ-invariant action of G on X. A set $A \in \mathfrak{B}$ is said to be $\mathcal{T}$-invariant $\mod \mu$ if and only if $\mu(T^{-g}A \triangle A) = 0$ holds for all $g \in G$, where $\triangle$ denotes the symmetric difference. We write $\mathcal{I}_\mu(\mathcal{T}) := \{A \in \mathfrak{B} : A \text{ is } \mathcal{T}\text{-invariant} \mod \mu\}$. If $\mu(A) = 0$ or $\mu(A) = 1$ for all $A \in \mathcal{I}_\mu(\mathcal{T})$, then the m.p.d.s. $(X, \mathfrak{B}, \mu, \mathcal{T})$ is said to be *ergodic*. A family of measurable sets $\alpha = \{A_i\}_{i \in I}$ is called a *μ-partition* of X if $\mu(A_i \cap A_j) = 0$ $(i \neq j)$, $\mu\left(X \setminus \bigcup_{i \in I} A_i\right) = 0$ and $\mu(A_i) > 0$ $(\forall i \in I)$. Let α be a μ-partition of X. The *information* of α is the function I_α on X defined by $I_\alpha(x) := -\sum_{A \in \alpha}(\log_2 \mu(A)) \cdot 1_A(x)$ $(\forall x \in X)$. The *entropy* of α is defined by the average information, i.e., $H_\mu(\alpha) := \int_X I_\alpha d\mu = \sum_{A \in \alpha} \varphi(\mu(A))$ where $\varphi(t) := -t \log_2 t$. From Kolmogorov complexity's point of view, we choose the binary logarithm $\log_2$ instead of $\log_e$. Let β be another μ-partition. The *common refinement* of α and β is $\alpha \vee \beta := \{A \cap B : A \in \alpha, B \in \beta, \mu(A \cap B) > 0\}$. We set $T^{-g}\alpha := \{T^{-g}A : A \in \alpha\}$ for each $g \in G$ and $\alpha^\Lambda :=$

$\bigvee_{g\in\Lambda} T^{-g}\alpha$ for a finite subset $\Lambda \subset G$. The *dynamical entropy* of the m.p.d.s. $(X, \mathfrak{B}, \mu, \mathcal{T})$ relative to the partition α is $h(\mu, \alpha, \mathcal{T}) := \inf_{n>0} \frac{1}{|\Lambda_n|} H_\mu(\alpha^{\Lambda_n}) = \lim_{n\to\infty} \frac{1}{|\Lambda_n|} H_\mu(\alpha^{\Lambda_n})$.

Theorem 2.1 (Shannon-McMillan-Breiman) *Let $(X, \mathfrak{B}, \mu, \mathcal{T})$ be an ergodic m.p.d.s. and α be a μ-partition of X with $H_\mu(\alpha) < \infty$. Then*

$$h(\mu, \alpha, \mathcal{T}) = \lim_{n\to\infty} \frac{1}{|\Lambda_n|} I_{\alpha^{\Lambda_n}} \quad \textit{in } L^1(X, \mu).$$

Moreover, if α is finite, then this convergence holds also for μ-a.s. $x \in X$.

The *Kolmogorov-Sinai entropy* of the m.p.d.s $(X, \mathfrak{B}, \mu, \mathcal{T})$ is defined by

$$h_{\mathcal{T}}(\mu) := \sup\{h(\mu, \alpha, \mathcal{T}) : \alpha \text{ is a } \mu\text{-partition with } H_\mu(\alpha) < \infty\}.$$

We denote by α^G the σ-algebra generated by all $T^{-g}\alpha,\ g \in G$. A μ-partition α is called a *μ-generator* if $\alpha^G = \mathfrak{B} \mod \mu$, where this equation means that $\forall A \in \mathfrak{B}, \exists B \in \alpha^G, \mu(A \triangle B) = 0$.

Theorem 2.2 (Kolmogorov-Sinai) *Let $(X, \mathfrak{B}, \mu, \mathcal{T})$ be a m.p.d.s. and α be a μ-generator such that $H(\alpha) < \infty$. Then $h_{\mathcal{T}}(\mu) = h(\mu, \alpha, \mathcal{T})$.*

Let $(X, \mathcal{T})$ be a *topological dynamical system* (t.d.s.), namely, X be a compact metrizable space and $\mathcal{T} = (T^g)_{g\in G}$ be a continuous action of G on X. In this setting we denote by $\mathfrak{B}(X)$ the Borel σ-algebra of X. We denote by $M(X)$ the set of all probability measures on the Borel measurable space $(X, \mathfrak{B}(X))$, by $M(X, \mathcal{T})$ the set of all $\mathcal{T}$-invariant probability measures on $(X, \mathfrak{B}(X))$ and by $EM(X, \mathcal{T})$ the set of all ergodic members in $M(X, \mathcal{T})$, respectively.

2.2 Kolmogorov complexity

Let Σ be a finite set and $|\Sigma| \geq 2$. Without loss of generality, we set $\Sigma := \{0, 1, \cdots, N\}$ where $N \in \mathbb{N}$. We define the set of all finite *strings* over Σ as

$$\Sigma^* := \bigcup_{n=0}^{\infty} \Sigma^n = \{\lambda, 0, 1, \cdots, N, 00, 01, \cdots, 0N, 10, \cdots, 1N, \cdots, NN, 000, \cdots\},$$

where $\Sigma^0 = \{\lambda\}$ and λ denote the empty string. The *length* of $x \in \Sigma^*$ is denoted by $l(x)$. For all $x, y \in \Sigma^*$, we call x a *prefix* of y if there exists $z \in \Sigma^*$ such that $y = xz$. A set $A \subset \Sigma^*$ is said to be *prefix-free* if, for all $x \in A$, the elements of $A \setminus \{x\}$ are not prefixes of x. Let $\mathcal{D}$ be a subset of $\{0,1\}^*$ and let f be a function from $\mathcal{D}$ to Σ^*. If $\mathcal{D} \subsetneq \{0,1\}^*$, we call such a function f a *partial function* and write $f : \{0,1\}^* \rightsquigarrow \Sigma^*$, and if $\mathcal{D} = \{0,1\}^*$ then we call f a *total function.* A partial function $\phi : \{0,1\}^* \rightsquigarrow \Sigma^*$ is said to be *partial recursive* if and only if there exists a Turing machine M such that ϕ is computed by M, i.e., for all $x \in \{0,1\}^*$, M on input x halts if and only if $x \in \mathrm{dom}(\phi)$, in that case, M

outputs $\phi(x)$. Moreover, if $\mathrm{dom}(\phi)$ is prefix-free, then we call ϕ a *partial recursive prefix function*. Let $\phi : \{0,1\}^* \rightsquigarrow \Sigma^*$ be a partial recursive prefix function. For all $x \in \Sigma^*$, the *complexity* of x with respect to ϕ is defined by

$$K_\phi(x) := \begin{cases} \min\{l(p) : p \in \phi^{-1}(x)\}, & (\phi^{-1}(x) \neq \emptyset), \\ \infty & (\phi^{-1}(x) = \emptyset). \end{cases}$$

A partial recursive prefix function $\phi : \{0,1\}^* \rightsquigarrow \Sigma^*$ is said to be *additively optimal* if for all partial recursive prefix functions $\psi : \{0,1\}^* \rightsquigarrow \Sigma^*$, there exists a constant $c_{\phi,\psi} \in \mathbb{R}$ such that for all $x \in \Sigma^*$, $K_\phi(x) \leq K_\psi(x) + c_{\phi,\psi}$. We fix such a function ϕ and define the *prefix Kolmogorov complexity* of $x \in \Sigma^*$ by $K(x) := K_\phi(x)$.

2.3 Shift dynamical system

Let $\Sigma := \{0, 1, \cdots, N\}$ ($N \in \mathbb{N}$) and we set $\Omega := \Sigma^G$. By Tychonoff's theorem, Ω endowed with the product topology of the discrete topology on Σ is a compact topological space. For all $n \in \mathbb{N}$ and for all $s \in \Sigma^{\Lambda_n}$, we define the *cylinder set* of s by $[\![s]\!] := \{\omega \in \Omega : \omega \restriction \Lambda_n = s\}$. We set

$$\Sigma^{\Lambda_*} := \bigcup_{n=0}^{\infty} \Sigma^{\Lambda_n}$$

where $\Sigma^{\Lambda_0} := \{\lambda\}$ and write $[\![V]\!] := \bigcup_{s \in V} [\![s]\!]$ for all $V \subset \Sigma^{\Lambda_*}$. Let $\sigma^g : \Omega \to \Omega$ denote the *shift* by $g \in G$, i.e., $(\sigma^g \omega)_i := \omega_{i+g}$ for all $\omega = (\omega_i)_{i \in G}$, and we write $\sigma := (\sigma^g)_{g \in G}$. Since σ is a continuous action of G on Ω, (Ω, σ) is a t.d.s.. Note that for all $\mu \in M(\Omega, \sigma)$, the partition $\{[\![s]\!]\}_{s \in \Sigma^{\Lambda_1}}$ is a μ-generator. A nonempty subset $S \subset \Omega$ is called a *subshift* if and only if S is shift-invariant and closed. If $S \subset \Omega$ is a subshift, then $(S, \sigma \restriction S)$ is a t.d.s., where $\sigma \restriction S := (\sigma^g \restriction S)_{g \in G}$. We fix an arbitrary bijective computable function $f : \mathbb{Z}_+ \to G$ such that for all $n \in \mathbb{N}$, $f(\{0, 1, \cdots, |\Lambda_n| - 1\}) = \Lambda_n$ and define $\mathcal{G} : \Sigma^{\Lambda_*} \to \Sigma^*$ as follows:

$$\mathcal{G}(s) := \begin{cases} s_{f(0)} \cdots s_{f(|\Lambda_n|-1)}, & s = (s_g)_{g \in \Lambda_n} \in \Sigma^{\Lambda_n} \ (n \in \mathbb{N}), \\ \lambda, & s = \lambda. \end{cases}$$

We define the prefix Kolmogorov complexity of $s \in \Sigma^{\Lambda_*}$ by $\mathsf{K}(s) := K(\mathcal{G}(s))$.

Lemma 2.3 *For all $n, k \in \mathbb{N}$, $|\{s \in \Sigma^{\Lambda_n} : \mathsf{K}(s) < k\}| < 2^k$.*

Proof. By [3, Theorem 7.2.4], we have for all $n, k \in \mathbb{N}$,

$$\begin{aligned} |\{s \in \Sigma^{\Lambda_n} : \mathsf{K}(s) < k\}| &= |\{\mathcal{G}(s) \in \Sigma^{|\Lambda_n|} : K(\mathcal{G}(s)) < k\}| \\ &\leq |\{x \in \Sigma^* : K(x) < k\}| < 2^k. \end{aligned}$$

□

The *upper* and *lower Kolmogorov complexity density* of $\omega \in \Omega$ are defined by

$$\overline{\mathcal{K}}(\omega) := \limsup_{n\to\infty} \frac{\mathsf{K}(\omega \restriction \Lambda_n)}{|\Lambda_n|}, \quad \underline{\mathcal{K}}(\omega) := \liminf_{n\to\infty} \frac{\mathsf{K}(\omega \restriction \Lambda_n)}{|\Lambda_n|}.$$

If $\overline{\mathcal{K}}(\omega) = \underline{\mathcal{K}}(\omega)$, we simply denote them by $\mathcal{K}(\omega)$. The quantities $\overline{\mathcal{K}}(\omega)$ and $\underline{\mathcal{K}}(\omega)$ are independent of the choice of additively optimal partial recursive prefix function ϕ and $\mathcal{G}$, and uniquely defined.

Lemma 2.4 *The functions $\overline{\mathcal{K}}, \underline{\mathcal{K}} : \Omega \to \mathbb{R}$ are measurable.*

Proof. Let us show that $\overline{\mathcal{K}}$ is measurable. For all $x \in \mathbb{R}$, we have

$$\begin{aligned}\overline{\mathcal{K}}^{-1}((-\infty, x)) &= \left\{\omega \in \Omega : \limsup_{n\to\infty} \frac{\mathsf{K}(\omega \upharpoonright \Lambda_n)}{|\Lambda_n|} < x\right\} \\ &= \bigcup_{k\in\mathbb{N}} \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \left\{\omega \in \Omega : \frac{\mathsf{K}(\omega \upharpoonright \Lambda_n)}{|\Lambda_n|} < x - \frac{1}{k}\right\}.\end{aligned}$$

Here

$$\left\{\omega : \frac{\mathsf{K}(\omega \upharpoonright \Lambda_n)}{|\Lambda_n|} < x - \frac{1}{k}\right\} = \begin{cases} \bigcup_{l=0}^{\lceil |\Lambda_n|\left(x-\frac{1}{k}\right)-1\rceil} \{\omega : \mathsf{K}(\omega \upharpoonright \Lambda_n) = l\}, & x - \frac{1}{k} > 0, \\ \emptyset, & x - \frac{1}{k} \leq 0. \end{cases}$$

Since the set $\{\omega \in \Omega : \mathsf{K}(\omega \upharpoonright \Lambda_n) = l\}$ is cylinder, then the set $\overline{\mathcal{K}}^{-1}((-\infty, x))$ is measurable. Hence the function $\overline{\mathcal{K}}$ is measurable. The proof for $\underline{\mathcal{K}}$ is similar. □

Remark 2.5 *Let C be a plain Kolmogorov complexity (that is not conditioned on prefix function). By [7, Example 3.1.4], we have for all $s \in \Sigma^{\Lambda_*}$*

$$C(\mathcal{G}(s)) \leq K(\mathcal{G}(s)) \leq C(\mathcal{G}(s)) + 2\log C(\mathcal{G}(s)).$$

It means that C and K are asymptotically equal. Then we may use C to define $\overline{\mathcal{K}}, \underline{\mathcal{K}}$.

3 Relation between KS entropy and Kolmogorov complexity

Let $d \in \mathbb{N}$, $G = \mathbb{Z}^d$ or $G = \mathbb{Z}^d_+$, $\Sigma = \{0, 1, \cdots, N\}$ $(N \in \mathbb{N})$ and $S \subset \Omega$ $(:= \Sigma^G)$ be a subshift. Other notations are the same as before. We set $\varsigma := \sigma \upharpoonright S$. Note that (S, ς) is a t.d.s.. We now state the main result.

Theorem 3.1 *If $\mu \in EM(S, \varsigma)$, then*

$$\mathcal{K}(\omega) = h_\varsigma(\mu), \quad \mu\text{-a.e.}\omega \in S. \tag{3.1}$$

Remark 3.2 *Brudno's original result is on the case $G = \mathbb{Z}_+$ only [2]. In the case $G = \mathbb{Z}^d$ or $G = \mathbb{Z}^d_+$, Simpson showed that if μ is a measure of maximal entropy, then (3.1) holds [10]. Our theorem is a generalization of them.*

It is sufficient to prove the theorem for the case $S = \Omega$. Because, if $\mu \in EM(\Omega, \sigma)$ and $\mu(S) = 1$, then $\mu \restriction S \in EM(S, \varsigma)$ and $h_\varsigma(\mu \restriction S) = h_\sigma(\mu)$ hold where $\mu \restriction S$ denotes the restriction of μ to S. So we prove the theorem about full shift (Ω, σ).

Theorem 3.3 (μ-typical sets) *Let $\mu \in EM(\Omega, \sigma)$. For all $\epsilon > 0$ and $n \in \mathbb{N}$, we set*

$$\mathfrak{T}_\epsilon^{(n)} := \left\{ s \in \Sigma^{\Lambda_n} : 2^{-|\Lambda_n|(h_\sigma(\mu)+\epsilon)} < \mu([\![s]\!]) < 2^{-|\Lambda_n|(h_\sigma(\mu)-\epsilon)} \right\},$$
$$\mathfrak{T}_\epsilon := \liminf_{n\to\infty} [\![\mathfrak{T}_\epsilon^{(n)}]\!].$$

Then the following holds:

$$\mu(\mathfrak{T}_\epsilon) = \lim_{n\to\infty} \mu([\![\mathfrak{T}_\epsilon^{(n)}]\!]) = 1 \quad \textit{and} \quad \limsup_{n\to\infty} \frac{\log_2 |\mathfrak{T}_\epsilon^{(n)}|}{|\Lambda_n|} \leq h_\sigma(\mu) + \epsilon.$$

Proof. It follows from Shannon-McMillan-Breiman theorem (Theorem 2.1). □

Lemma 3.4 *If $\mu \in EM(\Omega, \sigma)$, then*

$$\underline{\mathcal{K}}(\omega) \geq h_\sigma(\mu), \quad \mu\text{-}a.e.\omega \in \Omega. \tag{3.2}$$

Proof. If $h_\sigma(\mu) = 0$, then (3.2) is obvious. Let $h_\sigma(\mu) > 0$ and fix an arbitrary $k \in \mathbb{N}$ such that $\frac{1}{k} < h_\sigma(\mu)$. For all $n \in \mathbb{N}$, we set $D_{n,k} := \left\{ s \in \Sigma^{\Lambda_n} : \frac{\mathsf{K}(s)}{|\Lambda_n|} \leq h_\sigma(\mu) - \frac{1}{k} \right\}$. By Lemma 2.3, we have

$$|D_{n,k}| \leq 2^{|\Lambda_n|(h_\sigma(\mu) - \frac{1}{k}) + 1}. \tag{3.3}$$

We fix an arbitrary $\epsilon \in \left(0, \frac{1}{k}\right)$ and set $\mathfrak{T}^{(n)}_{\frac{1}{k}-\epsilon}$ as in Theorem 3.3. Then, by Theorem 3.3, we have for μ-a.e. $\omega \in \Omega$,

$$\exists N_\omega \in \mathbb{N}, \forall n \geq N_\omega, \omega \in [\![\mathfrak{T}^{(n)}_{\frac{1}{k}-\epsilon}]\!]. \tag{3.4}$$

On the other hand, by (3.3) and the definition of $\mathfrak{T}^{(n)}_{\frac{1}{k}-\epsilon}$, we have

$$\begin{aligned}\mu([\![D_{n,k}]\!] \cap [\![\mathfrak{T}^{(n)}_{\frac{1}{k}-\epsilon}]\!]) = \mu\left(\bigcup_{s \in D_{n,k} \cap \mathfrak{T}^{(n)}_{\frac{1}{k}-\epsilon}} [\![s]\!] \right) &\leq \sum_{s \in D_{n,k} \cap \mathfrak{T}^{(n)}_{\frac{1}{k}-\epsilon}} \mu([\![s]\!]) \\ &\leq 2^{|\Lambda_n|(h_\sigma(\mu) - \frac{1}{k}) + 1} \cdot 2^{-|\Lambda_n|(h_\sigma(\mu) - \frac{1}{k} + \epsilon)} = 2^{-|\Lambda_n|\epsilon + 1}.\end{aligned}$$

Hence $\sum_{n=1}^\infty \mu([\![D_{n,k}]\!] \cap [\![\mathfrak{T}^{(n)}_{\frac{1}{k}-\epsilon}]\!]) < \infty$ holds. Therefore, by the Borel-Cantelli lemma, for μ-a.e. $\omega \in \Omega$,

$$\exists N'_\omega \in \mathbb{N}, \forall n \geq N'_\omega, \omega \notin [\![D_{n,k}]\!] \cap [\![\mathfrak{T}^{(n)}_{\frac{1}{k}-\epsilon}]\!]. \tag{3.5}$$

By (3.4) and (3.5), for μ-a.e. $\omega \in \Omega$, we have

$$\exists N''_\omega \in \mathbb{N}, \forall n \geq N''_\omega, \omega \notin [\![D_{n,k}]\!].$$

Since $\omega \notin [\![D_{n,k}]\!]$ means $\frac{\mathsf{K}(\omega \upharpoonright \Lambda_n)}{|\Lambda_n|} > h_\sigma(\mu) - \frac{1}{k}$, we have for all $k \geq \lceil \frac{1}{h_\sigma(\mu)} \rceil + 1$,

$$\underline{\mathcal{K}}(\omega) \geq h_\sigma(\mu) - \frac{1}{k}, \quad \mu\text{-a.e.}\omega \in \Omega. \tag{3.6}$$

Hence

$$\begin{aligned}\mu\left(\{\omega : \underline{\mathcal{K}}(\omega) < h_\sigma(\mu)\}\right) &= \mu\left(\bigcup_{k=\lceil \frac{1}{h_\sigma(\mu)}\rceil+1}^{\infty} \left\{\omega : \underline{\mathcal{K}}(\omega) < h_\sigma(\mu) - \frac{1}{k}\right\}\right) \\ &\leq \sum_{k=\lceil \frac{1}{h_\sigma(\mu)}\rceil+1}^{\infty} \mu\left(\left\{\omega : \underline{\mathcal{K}}(\omega) < h_\sigma(\mu) - \frac{1}{k}\right\}\right) = 0.\end{aligned}$$

Therefore (3.2) holds. □

The following theorem plays a key role to prove the inverse direction.

Theorem 3.5 (Universally typical sets) *For all rational number $h_0 \in (0, \log_2 |\Sigma|]$, there exists a sequence of subsets $\{\mathfrak{U}^{(n)}_{h_0} \subset \Sigma^{\Lambda_n}\}_n$ such that the following conditions hold:*

(1) For all $\mu \in EM(\Omega, \sigma)$ with $h_\sigma(\mu) < h_0$,

$$\mu(\mathfrak{U}_{h_0}) = \lim_{n\to\infty} \mu([\![\mathfrak{U}^{(n)}_{h_0}]\!]) = 1 \quad \text{and} \quad \lim_{n\to\infty} \frac{\log_2 |\mathfrak{U}^{(n)}_{h_0}|}{|\Lambda_n|} = h_0$$

hold where $\mathfrak{U}_{h_0} := \liminf_{n\to\infty} [\![\mathfrak{U}^{(n)}_{h_0}]\!]$.

(2) The sequence of subsets $\{\mathfrak{U}^{(n)}_{h_0} \subset \Sigma^{\Lambda_n}\}_n$ is computable.

Proof. See [5, Thoerem 3.1] and its proof. □

Lemma 3.6 *If $\mu \in EM(\Omega, \sigma)$, then*

$$\overline{\mathcal{K}}(\omega) \leq h_\sigma(\mu), \quad \mu\text{-a.e.}\omega \in \Omega. \tag{3.7}$$

Proof. Let $n \in \mathbb{N}$ and $h_0 \in (h_\sigma(\mu), \log_2 |\Sigma|]$ be a rational number. We set $\mathfrak{U}^{(n)}_{h_0}$ and $\mathfrak{U}_{h_0}$ as in Theorem 3.5. By its definition, for all $\omega \in \mathfrak{U}_{h_0}$, we have

$$\exists N \in \mathbb{N},\ \forall n > N,\ \omega \in [\![\mathfrak{U}^{(n)}_{h_0}]\!]. \tag{3.8}$$

Note that $\omega \in [\![\mathfrak{U}_{h_0}^{(n)}]\!]$ means $\omega \upharpoonright \Lambda_n \in \mathfrak{U}_{h_0}^{(n)}$, and here we can encode each $s \in \mathfrak{U}_{h_0}^{(n)}$ into $\lceil \log_2 |\mathfrak{U}_{h_0}^{(n)}| \rceil$ bits code. Then the following holds for all $\omega \in \mathfrak{U}_{h_0}$:

$$\exists N \in \mathbb{N},\ \forall n > N,\ \frac{C(\mathcal{G}(\omega \upharpoonright \Lambda_n))}{|\Lambda_n|} \leq \frac{\log_2 |\mathfrak{U}_{h_0}^{(n)}| + \log_2 n + \text{const.}}{|\Lambda_n|}, \tag{3.9}$$

where C be a plain Kolmogorov complexity. As previously stated in Remark 2.5, C and K are asymptotically equal. Then by Theorem 3.5 and (3.9), we have $\overline{\mathcal{K}}(\omega) \leq h_0$ for all $\omega \in \mathfrak{U}_{h_0}$, namely, $\mathfrak{U}_{h_0} \subset \{\omega \in \Omega : \overline{\mathcal{K}}(\omega) \leq h_0\}$. Let $h_{0,k} \in (h_\sigma(\mu), \log_2 |\Sigma|]$ $(k \in \mathbb{N})$ be rational numbers with $\lim_{k\to\infty} h_{0,k} = h_\sigma(\mu)$. Then we have

$$\begin{aligned}\mu(\{\omega : \overline{\mathcal{K}}(\omega) > h_\sigma(\mu)\}) &= \mu\left(\bigcup_{k=1}^{\infty} \{\omega : \overline{\mathcal{K}}(\omega) > h_{0,k}\}\right) \\ &\leq \sum_{k=1}^{\infty} \mu\left(\{\omega : \overline{\mathcal{K}}(\omega) > h_{0,k}\}\right) \leq \sum_{k=1}^{\infty} \mu(\mathfrak{U}_{h_{0,k}}^c) = 0.\end{aligned}$$

Therefore (3.7) holds. □

Theorem 3.1 follows from Lemma 3.4 and Lemma 3.6.

Remark 3.7 *In Theorem 3.1, μ is not necessarily computable. Especially if μ is a computable measure, then Theorem 3.1 is easily seen by the following way: Let $\nu \in M(\Sigma^{\mathbb{Z}_+})$ be a computable measure such that for all $\omega \in \Omega, \mu([\![\omega \upharpoonright \Lambda_n]\!]) = \nu([\![\mathcal{G}(\omega \upharpoonright \Lambda_n)]\!])$. By [6, THEOREM 5.1, LEMMA 5.2], if $\mathcal{G}(\omega)(:= \lim_{n\to\infty} \mathcal{G}(\omega \upharpoonright \Lambda_n), \omega \in \Omega)$ is Martin-Löff random with respect to ν, then there exist $c_1, c_2 > 0$ such that for all $n \in \mathbb{N}$*

$$\begin{aligned}-\log_2 \nu([\![\mathcal{G}(\omega \upharpoonright \Lambda_n)]\!]) - c_1 &< K(\mathcal{G}(\omega \upharpoonright \Lambda_n)) \\ &\leq -\log_2 \nu([\![\mathcal{G}(\omega \upharpoonright \Lambda_n)]\!]) + 2\log_2 l(\mathcal{G}(\omega \upharpoonright \Lambda_n)) + c_2.\end{aligned}$$

Then we have for μ-a.e. $\omega \in \Omega$

$$\mathcal{K}(\omega) = \lim_{n\to\infty} \frac{-\log_2 \mu([\![\omega \upharpoonright \Lambda_n]\!])}{|\Lambda_n|} = h_\sigma(\mu).$$

The last equality is derived from Shannon-McMillan-Breiman theorem.

Example 3.8 (d-dimensional Bernoulli shifts) *Let (Ω, σ) be the $\mathbb{Z}^d$ or $\mathbb{Z}_+^d$ shift space as before. We fix a probability vector $q = (q_i : i \in \Sigma)$ on Σ and denote the corresponding Bernoulli measure on $\mathfrak{B}(\Omega)$ by $\mu := q^{\times G}$. Then, by Kolmogorov-Sinai theorem (Theorem 2.2), we can show that $h_\sigma(\mu) = \sum_{i\in\Sigma} \varphi(q_i)$. By Theorem 3.1, we have for μ-a.e. $\omega \in \Omega$*

$$\mathcal{K}(\omega) = \sum_{i\in\Sigma} \varphi(q_i).$$

Corollary 3.9 *If $\mu \in M(\Omega,\sigma)$, then there exists $\mathcal{K}(\omega) = \lim_{n\to\infty} \frac{\mathsf{K}(\omega\restriction\Lambda_n)}{|\Lambda_n|}$ for μ-a.e. $\omega \in \Omega$ and the following holds:*

$$h_\sigma(\mu) = \mu(\mathcal{K}) = \lim_{n\to\infty} \frac{1}{|\Lambda_n|} \sum_{s\in\Sigma^{\Lambda_n}} \mathsf{K}(s)\mu([\![s]\!]).$$

Proof. Let $\mu = \int_{EM(\Omega,\sigma)} \nu d\rho(\nu)$ be the ergodic decomposition, where ρ be a probability measure on $EM(\Omega,\sigma)$ (see [4, 9, 11]). By Jacobs's theorem [11, Theorem 8.4] and Theorem 3.1, we have

$$\mu(\overline{\mathcal{K}}) = \int_{EM(\Omega,\sigma)} \left\{ \int_\Omega \overline{\mathcal{K}}(\omega)d\nu(\omega) \right\} d\rho(\nu) = \int_{EM(\Omega,\sigma)} h_\sigma(\nu)d\rho(\nu) = h_\sigma(\mu)$$

and $\mu(\underline{\mathcal{K}}) = h_\sigma(\mu)$ is also the same. Hence for μ-a.e. $\omega \in \Omega$ there exists $\mathcal{K}(\omega)$ and $\mu(\mathcal{K}) = h_\sigma(\mu)$ holds. On the other hand, by (3.9) and Lebesgue's convergence theorem, we have

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{|\Lambda_n|} \sum_{s\in\Sigma^{\Lambda_n}} \mathsf{K}(s)\mu([\![s]\!]) &= \lim_{n\to\infty} \int_\Omega \frac{\mathsf{K}(\omega\restriction\Lambda_n)}{|\Lambda_n|} d\mu(\omega) \\ &= \int_\Omega \lim_{n\to\infty} \frac{\mathsf{K}(\omega\restriction\Lambda_n)}{|\Lambda_n|} d\mu(\omega) = \mu(\mathcal{K}).\end{aligned}$$

□

Remark 3.10 *In the case $G = \mathbb{Z}_+$, Corollary 3.9 can be found in [1].*

Example 3.11 (d-dimensional Ising model) *Let $d \in \mathbb{N}$ and $\Sigma := \{+1, -1\}$. Here $+1$ and -1 represent "spin up" and "spin down" at the sites of a "lattice gas" on $G := \mathbb{Z}^d$, respectively. Let $\Omega := \Sigma^G$ be a configuration space and σ be a shift action of G on Ω. For d-dimensional Ising model, the local energy function $\psi : \Omega \to \mathbb{R}$ is defined by*

$$\psi(\omega) := -\beta \left(-\sum_{j=1}^{d} (\omega_{\mathbf{0}}\omega_{e_j} + \omega_{\mathbf{0}}\omega_{-e_j}) - B\omega_{\mathbf{0}} \right), \quad \omega \in \Omega,$$

where $\mathbf{0} := (0, \cdots, 0), e_j := (0, \cdots, \overset{j\text{th}}{1}, \cdots, 0) \in G$. Here $-\sum_{j=1}^{d}(\omega_{\mathbf{0}}\omega_{e_j} + \omega_{\mathbf{0}}\omega_{-e_j})$ represents the interaction between neighboring spins, $-B\omega_{\mathbf{0}}$ represents the effect of a magnetic field $B \in \mathbb{R}$ on the spin at site $\mathbf{0}$ and $\beta \geq 0$ denote the inverse temperature. Then the pressure of this model is given by

$$p(\psi) = \sup_{\mu\in M(\Omega,\sigma)} \mu(\mathcal{K} + \psi).$$

In mathematical point of view, this example is just a replacement of $h_\sigma(\mu)$ by $\mu(\mathcal{K})$, but it shows that the generalization of Brudno's theorem to $\mathbb{Z}^d$-action (especially $d = 2$ or 3) has a physical background.

By using Brudno's theorem for multidimensional subshifts, we can construct a universally typical sets of multidimensional data as follows.

Theorem 3.12 (Universally typical sets using Brudno's theorem) *Let $h_0 > 0$ and $n \in \mathbb{N}$. We set*

$$\mathfrak{K}_{h_0}^{(n)} = \left\{ s \in \Sigma^{\Lambda_n} : \frac{\mathsf{K}(s)}{|\Lambda_n|} < h_0 \right\} \quad \textit{and} \quad \mathfrak{K}_{h_0} := \liminf_{n\to\infty} [\![\mathfrak{K}_{h_0}^{(n)}]\!].$$

Then for all $\mu \in EM(\Omega, \sigma)$ with $h_\sigma(\mu) < h_0$ the following holds:

$$\mu(\mathfrak{K}_{h_0}) = \lim_{n\to\infty} \mu([\![\mathfrak{K}_{h_0}^{(n)}]\!]) = 1 \quad \textit{and} \quad \limsup_{n\to\infty} \frac{\log_2 |\mathfrak{K}_{h_0}^{(n)}|}{|\Lambda_n|} \le h_0.$$

Proof. For all $\mu \in EM(\Omega, \sigma)$ with $h_\sigma(\mu) < h_0$, by Theorem 3.1, we have

$$\begin{aligned}
1 &= \mu\left(\left\{ \omega \in \Omega : \lim_{n\to\infty} \left| \frac{\mathsf{K}(\omega \upharpoonright \Lambda_n)}{|\Lambda_n|} - h_\sigma(\mu) \right| = 0 \right\} \right) \\
&= \mu\left(\bigcap_{\epsilon>0} \bigcup_{N\in\mathbb{N}} \bigcap_{n>N} \left\{ \omega \in \Omega : h_\sigma(\mu) - \epsilon < \frac{\mathsf{K}(\omega \upharpoonright \Lambda_n)}{|\Lambda_n|} < h_\sigma(\mu) + \epsilon \right\} \right) \\
&\le \mu\left(\bigcup_{N\in\mathbb{N}} \bigcap_{n>N} \left\{ \omega \in \Omega : \frac{\mathsf{K}(\omega \upharpoonright \Lambda_n)}{|\Lambda_n|} < h_0 \right\} \right) = \mu(\mathfrak{K}_{h_0}) \le \liminf_{n\to\infty} \mu([\![\mathfrak{K}_{h_0}^{(n)}]\!]).
\end{aligned}$$

Then $\mu(\mathfrak{K}_{h_0}) = \lim_{n\to\infty} \mu([\![\mathfrak{K}_{h_0}^{(n)}]\!]) = 1$ holds. Since $\lim_{n\to\infty} \mu([\![\mathfrak{K}_{h_0}^{(n)}]\!]) = 1$, $|\mathfrak{K}_{h_0}^{(n)}| \neq 0$ holds for sufficiently large $n \in \mathbb{N}$. Therefore, by Lemma 2.3, we have $0 \neq |\mathfrak{K}_{h_0}^{(n)}| < 2^{h_0|\Lambda_n|+1}$ ($n \gg 1$). Hence $\frac{\log_2 |\mathfrak{K}_{h_0}^{(n)}|}{|\Lambda_n|} < h_0 + \frac{1}{|\Lambda_n|}$ ($n \gg 1$). This completes the proof. □

Acknowledgments

We are grateful to A. Arai for reading our paper and for his valuable comments. We also thank K. Tadaki for his important comments about how we should define the prefix Kolmogorov complexity of $s \in \Sigma^{\Lambda_*}$. We would also like to express our sincere gratitude to M. Yuri for her valuable comments and longstanding encouragements. T. Fuda would like to thank S. Galatolo and A. Sakai for valuable discussions. We thank M. Kuroda for his cooperation. Finally, we would like to thank two anonymous referees for their constructive comments which improved the paper.

References

[1] V. Benci, C. Bonanno, S. Galatolo, G. Menconi, M. Virgilio, Dynamical systems and computable information, Discrete Contin. Dyn. Syst., Ser. B **4**, 935–960 (2004)

[2] A. A. Brudno, Entropy and the complexity of the trajectories of a dynamical system, Trans. Mosc. Math. Soc. **2**, 127–151 (1983)

[3] T. M. Cover, J. A. Thomas, Elements of information theory, John Wiley & Sons (1991)

[4] G. Keller, Equilibrium States in Ergodic Theory, Cambridge University Press (1998)

[5] T. Krüger, G. Montufar, R. Seiler, R. Siegmund-Schultze, Universally typical sets for ergodic sources of multidimensional data, Kybernetika **49**, 868–882 (2013)

[6] M. Van Lambalgen, Von Mises' definition of random sequences reconsidered, J. Symb. Logic. **52**, 725–755 (1987)

[7] M. Li, P. Vitányi, An Introduction to Kolmogorov Complexity and Its Applications, 3rd edn., Springer (2008)

[8] D. Ornstein, B. Weiss, The Shannon-McMillan-Breiman theorem for a class of amenable groups, Israel J. Math. **44**, 53–60 (1983)

[9] M. Pollicott, M. Yuri, Dynamical Systems and Ergodic Theory, Cambridge University Press (1998)

[10] S. G. Simpson, Symbolic Dynamics: Entropy = Dimension = Complexity, Theory Comput. Syst. **56**, 527–543 (2015)

[11] P. Walters, An Introduction to Ergodic Theory, Springer (1982)