Pre-image Variational Principle for Bundle Random Dynamical Systems

Xianfeng Ma\textsuperscript{a,b}, Ercai Chen\textsuperscript{b,c}

\textsuperscript{a}Department of Mathematics, East China University of Science and Technology
Shanghai 200237, China
\textsuperscript{b}School of Mathematics and Computer Science, Nanjing Normal University
Nanjing 210097, China
\textsuperscript{c}Center of Nonlinear Science, Nanjing University
Nanjing 210093, China

Abstract

The pre-image topological pressure is defined for bundle random dynamical systems. A variational principle for it has also been given.

Key words: Random dynamical systems, Variational principle, Pre-image topological pressure

2000 MSC: 37D35, 37A35, 37H99

1. Introduction

In deterministic dynamical systems, the thermodynamic formalism based on the notions of pressure, of Gibbs and equilibrium states plays a fundamental role in statistic mechanics, ergodic theory and dimension theory \cite{3,13,18,23}. The background for the study of equilibrium states is an appropriate form of the variational principle. Its first version was formulated by Ruelle \cite{2}. In random dynamical systems (RDS), the thermodynamic formalism is also important in the study of chaotic properties of random transformations \cite{5,8,20,25}. The first version of the variational principle for random transformations was given by Ledrappier and Walters in the framework of the relativized ergodic theory \cite{6}, and it was extended by Bo-

\textit{Email addresses: xianfengma@gmail.com (Xianfeng Ma), ecchen@njnu.edu.cn (Ercai Chen)}
genschütz [17] to random transformations acting on one place. Later Kifer [26] gave the variational principle for random bundle transformations.

In recent years, the pre-image structure of a map has also been studied by many authors [4, 11, 14, 15, 21, 22, 28]. In deterministic dynamical systems, Fiebig [4] studied the relation between the classical topological entropy and the dispersion of pre-images. Cheng and Newhouse [21] introduced the notions of the pre-image entropies and obtained a variational principle which is similar to the standard one. Zeng [7] defined the notion of the pre-image pressure and investigated its relationship with invariant measures. He also established a variational principle for the pre-image pressure, which was a generalization of Cheng’s result for the pre-image entropy. In random dynamical systems, Zhu [27] introduced the analogous notions as that in the deterministic case and gave the analogs of many known results for entropies, such as Shannon-McMillan-Breiman Theorem, the Kolmogorov-Sinai Theorem, the Abromov-Rokhlin formula and the variational principle.

In this paper, we present the notion of the pre-image topological pressure and derive the corresponding variational principle for bundle random dynamical systems. In fact, we formulate a random variational principle between the pre-image topological pressure, the pre-image measure-theoretic entropy and some functions of the invariant measure. We also introduce a revised definition of the random pre-image topological entropy without any additional assumptions, while the original notion defined in [27] need a strong measurability condition. All results in [27] still hold for our new notion. For the probability space consisting of a single point, we establish the pre-image variational principle on any compact invariant subset for deterministic dynamical systems, which is a generalization of Zeng’s result [7] for the whole space. The method we use is in the framework of Misiurewicz’s elegant proof [10]. Kifer’s method [24] and Cheng’s technique [21] are also adopted in the argument of our theorem. In fact, our proof generalizes Kifer’s proof of the standard variational principle for random bundle transformations.

This paper is organized as follows. In Section 2, we define the pre-image measure-theoretic entropy as a conditional entropy of the induced skew product transformation and give another fiberwise expression for bundle random dynamical systems. In Section 3, we define the pre-image topological pressure for bundle random dynamical systems and give the power rule for this pressure. In Section 4, we state and prove the pre-image variational principle.
2. Pre-image measure-theoretic entropy for bundle RDS

Let \((\Omega, \mathcal{F}, \mathcal{P})\) be a probability space together with an invertible \(\mathcal{P}\)-preserving transformation \(\vartheta\), where \(\mathcal{F}\) is complete, countably generated and separated points. Let \((X, d)\) be a compact metric space together with the Borel \(\sigma\)-algebra \(\mathcal{B}\). Let \(\mathcal{E} \subset \Omega \times X\) be measurable with respect to the product \(\sigma\)-algebra \(\mathcal{F} \times \mathcal{B}\) and the fibers \(\mathcal{E}_\omega = \{ x \in X : (\omega, x) \in \mathcal{E} \}, \omega \in \Omega\) be compact. A continuous bundle random dynamical system (RDS) \(T\) over \((\Omega, \mathcal{F}, \mathcal{P}, \vartheta)\) is generated by map \(T_\omega : \mathcal{E}_\omega \to \mathcal{E}_{\vartheta \omega}\) with iterates \(T^n_\omega = T_{\vartheta^{n-1} \omega} \cdot \cdots \cdot T_{\vartheta \omega} T_\omega, n \geq 1\), and \(T^0_\omega = id\), so that the map \((\omega, x) \to T_\omega x\) is measurable and the map \(x \to T_\omega x\) is continuous for \(\mathcal{P}\)-almost all \((a.a.)\) \(\omega\). The map \(\Theta : \mathcal{E} \to \mathcal{E}\) defined by \(\Theta(\omega, x) = (\vartheta \omega, T_\omega x)\) is called the skew product transformation.

Let \(\mathcal{P}_\mathcal{P}(\mathcal{E}) = \{ \mu \in \mathcal{P}_\mathcal{P}(\Omega \times X) : \mu(\mathcal{E}) = 1 \}\), where \(\mathcal{P}_\mathcal{P}(\Omega \times X)\) is the space of probability measures on \(\Omega \times X\) with the marginal \(\mathcal{P}\) on \(\Omega\). Any \(\mu \in \mathcal{P}_\mathcal{P}(\mathcal{E})\) on \(\mathcal{E}\) can be disintegrated as \(d\mu(\omega, x) = d\mu_\omega(x) d\mathcal{P}(\omega)\) (See [16]), where \(\mu_\omega\) are regular conditional probabilities with respect to the \(\sigma\)-algebra \(\mathcal{F}_\mathcal{E}\) formed by all sets \((A \times X) \cap \mathcal{E}\) with \(A \in \mathcal{F}\). Let \(\mathcal{M}_\mathcal{P}(\mathcal{E}, T)\) be the set of \(\Theta\)-invariant measures \(\mu \in \mathcal{P}_\mathcal{P}(\mathcal{E})\). \(\mu\) is \(\Theta\)-invariant if and only if the disintegrations \(\mu_\omega\) of \(\mu\) satisfy \(T_\omega \mu_\omega = \mu_{\vartheta \omega}\) \(\mathcal{P}\)-a.s. [9]. Let \(\mathcal{Q} = \{ \mathcal{Q}_i \}\) be a finite measurable partition of \(\mathcal{E}\), and \(\mathcal{Q}(\omega) = \{ x \in \mathcal{E}_\omega : (\omega, x) \in \mathcal{Q}_i \}\) is a partition of \(\mathcal{E}_\omega\). For each \(\omega \in \Omega\), let \(\mathcal{B}_\omega = \{ B \cap \mathcal{E}_\omega : B \in \mathcal{B} \}\) and \(\mathcal{B}^- = \bigcap_{n \geq 0} (T^{-1}_\omega)^{-1} B_{\vartheta^n \omega}\). Similarly, let \((\mathcal{F} \times \mathcal{B})_{\mathcal{E}} = \{ C \cap \mathcal{E} : C \in \mathcal{F} \times \mathcal{B} \}\) and \((\mathcal{F} \times \mathcal{B})_{\mathcal{E}}^- = \bigcap_{n \geq 0} \Theta^{-n}(\mathcal{F} \times \mathcal{B})_{\mathcal{E}}\).

For \(\mu \in \mathcal{P}_\mathcal{P}(\mathcal{E})\), the conditional entropy of \(\mathcal{Q}\) given by the \(\sigma\)-algebra of \(\mathcal{F}_{\mathcal{E}} \cap (\mathcal{F} \times \mathcal{B})_{\mathcal{E}}\) is defined as usual (See Kifer [24]) by

\[
H_{\mu}(\mathcal{Q} \mid \mathcal{F}_{\mathcal{E}} \cap (\mathcal{F} \times \mathcal{B})_{\mathcal{E}}) = - \sum_i \mu(\mathcal{Q}_i \mid \mathcal{F}_{\mathcal{E}} \cap (\mathcal{F} \times \mathcal{B})_{\mathcal{E}}) \log \mu(\mathcal{Q}_i \mid \mathcal{F}_{\mathcal{E}} \cap (\mathcal{F} \times \mathcal{B})_{\mathcal{E}}) \, d\mu.
\]

The pre-image measure-theoretic entropy \(h^{(r)}_{\text{pre}, \mu}(T)\) of bundle RDS \(T\) with respect to \(\mu\) is defined by the formula

\[
h^{(r)}_{\text{pre}, \mu}(T) = \sup_{\mathcal{Q}} h^{(r)}_{\text{pre}, \mu}(T, \mathcal{Q}),
\]

where

\[
h^{(r)}_{\text{pre}, \mu}(T, \mathcal{Q}) = \lim_{n \to \infty} \frac{1}{n} H_{\mu}\left(\bigvee_{i=0}^{n-1} \Theta^{-i} \mathcal{Q} \mid \mathcal{F}_{\mathcal{E}} \cap (\mathcal{F} \times \mathcal{B})_{\mathcal{E}}\right).
\]
and the supremum is taken over all finite or countable measurable partitions \( \mathcal{Q} = \{\mathcal{Q}_i\} \) of \( \mathcal{E} \) with finite conditional entropy \( H_\mu(\mathcal{Q} \mid \mathcal{F}_\mathcal{E}) < \infty \). The existence of the limit follows from the formula \( \Theta^{-1}(\mathcal{F}_\mathcal{E} \vee (\mathcal{F} \times \mathcal{B})_\mathcal{E}) \subset \mathcal{F}_\mathcal{E} \vee (\mathcal{F} \times \mathcal{B})_\mathcal{E} \) and the standard subadditive argument (cf. Kifer [24, Theorem II.1.1]). The resulting entropy remains the same by taking the supremum only over partitions \( \mathcal{Q} \) of \( \mathcal{E} \) into sets \( Q_i \) of the form \( Q_i = (\Omega \times P_i) \cap \mathcal{E} \), where \( \mathcal{P} = \{P_i\} \) is a partition of \( X \) into measurable sets, so that \( Q_i(\omega) = P_i \cap \mathcal{E}_\omega \) (See [17, 19, 24] for detail).

Compared with Zhu [27], we define the pre-image measure-theoretic entropy on the measurable subset \( \mathcal{E} \) instead of on the whole space \( \Omega \times X \). Moreover, if \( \mathcal{E} = \Omega \times X \), then the above definition is just the measure-theoretic pre-image entropy in [27]. In this sense, the definition is a generalization of Zhu’s.

In [27], Zhu gave another fiberwise expression for his defined measure-theoretic pre-image entropy, which can be seen as a generalization of Kifer’s discussion on the standard measure-theoretic entropy for random bundle transformations [26]. In a similar way, we can give a fiberwise expression for the above definition. For completeness of this paper, we state this proposition and give the proof.

**Proposition 1.** Let \( \mathcal{Q} \) be a finite measurable partition of \( \mathcal{E} \). Then

\[
h^{(r)}_{\text{pre}, \mu}(T, \mathcal{Q}) = \lim_{n \to \infty} \frac{1}{n} \int H_\mu \left( \bigvee_{i=0}^{n-1} (T^{i}_\omega)^{-1} \mathcal{Q}(\cdot|B^{-}_\omega) \right) d\mathcal{P}(\omega). \tag{1}
\]

**Proof.** Note that for any \( f \in L^1(\Omega \times X, \mu) \) and \( \bigcup_{\omega \in F} (\omega) \times B_\omega \in \mathcal{F}_\mathcal{E} \vee (\mathcal{F} \times \mathcal{B})_\mathcal{E} \) where \( F \in \text{Pr}_\Omega \mathcal{F}_\mathcal{E} \) and \( B_\omega \in \mathcal{B}_\mathcal{E} \), \( \text{Pr}_\Omega \mathcal{F}_\mathcal{E} \) is the projection of \( \mathcal{F}_\mathcal{E} \) into \( \Omega \), we have

\[
\int_F \int_{B_\omega} E(f|\mathcal{F}_\mathcal{E} \vee (\mathcal{F} \times \mathcal{B})_\mathcal{E})(\omega, x) \, d\mu_\omega(x) \, d\mathcal{P}(\omega)
= \int_{\bigcup_{\omega \in F} (\omega) \times B_\omega} f(\omega, x) \, d\mu(\omega, x)
= \int_F \int_{B_\omega} f_\omega(x) \, d\mu_\omega(x) \, d\mathcal{P}(\omega)
= \int_F \int_{B_\omega} E(f_\omega|\mathcal{B}_\mathcal{E})(\omega) \, d\mu_\omega(x) \, d\mathcal{P}(\omega)
\]
where \( f_\omega(x) = f(\omega, x) \). Therefore,

\[
E(f|\mathcal{F}_E \vee (\mathcal{F} \times \mathcal{B})_E)(\omega, x) = E(f_\omega|\mathcal{B}_E^\omega)(x) \quad \mu - \text{a.e.}
\]

Hence for any finite measurable partition \( Q \) of \( \mathcal{E} \), we have

\[
I_\mu(Q|\mathcal{F}_E \vee (\mathcal{F} \times \mathcal{B})_E)(\omega, x) = I_\mu_\omega(Q(\omega)|\mathcal{B}_E^\omega)(x) \quad \mu - \text{a.e.}
\]

where \( I.(\cdot|\cdot) \) denotes the standard conditional information function. Thus

\[
H_\mu(Q|\mathcal{F}_E \vee (\mathcal{F} \times \mathcal{B})_E) = \int I_\mu_\omega(Q(\omega)|\mathcal{B}_E^\omega) \, dP(\omega).
\]

Since \((\Theta^{-i}Q)(\omega) = (T_\omega^i)^{-1}Q(\vartheta^i\omega)\) for any \( i \in \mathbb{N} \), then

\[
H_\mu\left( \bigvee_{i=0}^{n-1} \Theta^{-i}Q|\mathcal{F}_E \vee (\mathcal{F} \times \mathcal{B})_E \right) = \int H_\mu_\omega\left( \bigvee_{i=0}^{n-1} (T_\omega^i)^{-1}Q(\vartheta^i\omega)|\mathcal{B}_E^\omega \right) \, dP(\omega).
\]

Dividing by \( n \) and letting \( n \to \infty \), we obtain the equality (1). \( \Box \)

Moreover, if we use Zhu’s methods and restrict the whole space \( \Omega \times X \) to the measurable subset \( \mathcal{E} \) in [27], then all results with respect to the pre-image measure-theoretic entropy defined by Zhu also hold for the above definition. Then, we can use those results directly without giving any proof whenever we consider the pre-image measure-theoretic entropy.

### 3. Pre-image topological pressure for bundle RDS

Let \( X_\mathcal{E} = \{ x \in X : (\omega, x) \in \mathcal{E}, \omega \in \Omega \} \). For each \( x \in X_\mathcal{E} \), by the measurability of bundle RDS \( T \), \( \mathcal{E}(x) = \{ (\omega, y) : \omega \in \Omega, y \in T_{\omega}^{-1}x \} \) is measurable with respect to the product \( \sigma \)-algebra \( \mathcal{F} \times \mathcal{B} \). For each \( k \in \mathbb{N} \), let \( \mathcal{E}(x, k) = \{ (\omega, y) : \omega \in \Omega, y \in (T_{\omega}^k)^{-1}x \} \). By the continuity of bundle RDS \( T \) and [1, Theorem III.30], it is not hard to see that \( \mathcal{E}(x, k) \) is also measurable. Since for each \( k \in \mathbb{N}, (T_{\omega}^k)^{-1}x \) is compact in \( \mathcal{E}_\omega \), then the mapping \( \omega \mapsto (T_{\omega}^k)^{-1}x \) is measurable (See[1, Chaper III]) with respect to the Borel \( \sigma \)-algebra induced by the Hausdorff topology on the space \( K(X) \), and the distance function \( d(z, (T_{\omega}^k)^{-1}x) \) is measurable in \( \omega \in \Omega \) for each \( z \in X \).

For each \( n \in \mathbb{N} \), a family of metrics \( d_n^\omega \) on \( \mathcal{E}_\omega \) is defined as

\[
d_n^\omega(y, z) = \max_{0 \leq i < n} (d(T_{\omega}^iy, T_{\omega}^iz)), \quad y, z \in \mathcal{E}_\omega.
\]
It is not hard to see that for each \( k \in \mathbb{N} \) and \( x \in X_\mathcal{E} \), the set \( \mathcal{E}^{(2)}(x, k) = \{ (\omega, y, z) : y, z \in (T^k_\omega)^{-1}x \} \) belongs to the product \( \sigma \)-algebra \( \mathcal{F} \times \mathcal{B}^3 \) (See [1, Proposition III.13]). Since for each \( m \in \mathbb{N} \), \( \epsilon > 0 \) and a real number \( a \) the set \( \{ (\omega, y, z) \in \mathcal{E}^{(2)}(x, k) : d(T^m_\omega y, T^m_\omega z) \leq a \epsilon \} \) is measurable with respect to this product \( \sigma \)-algebra, then \( d_n^x(y, z) \) depends measurably on \( (\omega, y, z) \in \mathcal{E}^{(2)}(x, k) \).

For each \( n \in \mathbb{N} \) and \( \epsilon > 0 \), a set \( F \subset \mathcal{E}_\omega \) is said to be \((\omega, n, \epsilon)\)-separated if for any \( y, z \in F \), \( y \neq z \) implies \( d_n^x(y, z) > \epsilon \). Similarly, for a compact subset \( K \subset \mathcal{E}_\omega \), \( F \subset K \) is said to be \((\omega, n, \epsilon)\)-separated for \( K \) if for any \( y, z \in F \), \( y \neq z \) implies \( d_n^x(y, z) > \epsilon \).

Due to the compactness, there exists a smallest natural number \( s_n(\omega, \epsilon) \) such that \( \text{card}(F) \leq s_n(\omega, \epsilon) < \infty \) for every \((\omega, n, \epsilon)\)-separated \( F \). Moreover, there always exists a maximal \((\omega, n, \epsilon)\)-separated set \( F \) in the sense that for every \( y \in \mathcal{E}_\omega \) with \( y \not\in F \) the set \( F \cup \{ y \} \) is not \((\omega, n, \epsilon)\)-separated anymore. In particular, this is also true for any compact subset \( K \) of \( \mathcal{E}_\omega \). Let \( s_n(\omega, \epsilon, K) \) be the smallest natural number such that \( \text{card}(F) \leq s_n(\omega, \epsilon, K) < \infty \) for every \((\omega, n, \epsilon)\)-separated set \( F \) of \( K \).

For each function \( f \) on \( \mathcal{E} \), which is measurable in \((\omega, x)\) and continuous in \( x \in \mathcal{E}_\omega \), let

\[
\| f \| = \int \| f(\omega) \|_\infty \, d\mathbf{P}, \quad \text{where} \quad \| f(\omega) \|_\infty = \sup_{x \in \mathcal{E}_\omega} | f(\omega, x) | .
\]

Let \( L^1_\mathcal{E}(\Omega, \mathcal{C}(X)) \) be the space of such functions \( f \) with \( \| f \| < \infty \). If we identify \( f \) and \( g \) for \( f, g \in L^1_\mathcal{E}(\Omega, \mathcal{C}(X)) \) with \( \| f - g \| = 0 \), then \( L^1_\mathcal{E}(\Omega, \mathcal{C}(X)) \) is a Banach space with the norm \( \| \cdot \| \).

For any \( \delta > 0 \), let

\[
\kappa^{(f)}_\delta(\omega) = \sup \{ | f(\omega, x) - f(\omega, y) | : x, y \in \mathcal{E}_\omega, d(x, y) \leq \delta \}.
\]

For \( f \in L^1_\mathcal{E}(\Omega, \mathcal{C}(X)) \), \( k, n \in \mathbb{N} \) with \( k \geq n, \epsilon > 0 \), \( x \in X_\mathcal{E} \) with \((T^k_\omega)^{-1}x \neq \emptyset \), and an \((\omega, n, \epsilon)\)-separated set \( E \) of \((T^k_\omega)^{-1}x \) such that \( E \subset (T^k_\omega)^{-1}x \subset \mathcal{E}_\omega \), set

\[
S_n f(\omega, y) = \sum_{i=0}^{n-1} f(\Theta^i(\omega, y), T^i_\omega y) = \sum_{i=0}^{n-1} f \circ \Theta^i(\omega, y),
\]

and denote

\[
P_{\text{pre}, n, \omega}(T, f, \epsilon, (T^k_\omega)^{-1}x) = \sup_{y \in E} \sum_{y \in E} \exp S_n f(\omega, y),
\]

6
where the supremum is taken over all \((\omega, n, \epsilon)\)-separated sets of \((T_k^{(\omega)})^{-1}x\) in \(E_\omega\). Based on the foregoing analysis, clearly, any \((\omega, n, \epsilon)\)-separated set can be completed to a maximal one. Then, the supremum can be taken only over all maximal \((\omega, n, \epsilon)\)-separated sets. For \((T_k^{(\omega)})^{-1}x = \emptyset\), let \(P_{\text{pre}, n, \omega}(T, f, \epsilon, (T_k^{(\omega)})^{-1}x) = 0\) for all \(n\) and \(\epsilon\), then the function \(P_{\text{pre}, n, \omega}(T, f, \epsilon, (T_k^{(\omega)})^{-1}x)\) is well-defined. In this paper, we always assume that for each \(k \in N\) and \(x \in X_\epsilon\), \((T_k^{(\omega)})^{-1}x \neq \emptyset\).

Alternatively, we can also assume that for each \(\omega\) the mapping \(T^{(\omega)}\) is surjective.

The following auxiliary result, which relies on Kifer’s work \([26]\) and restricts his result to the family of compact subsets \((T_k^{(\omega)})^{-1}x\) of \(E_\omega\) for nonrandom positive number \(\epsilon\), provides the basic properties of measurability needed in what follows. We make a little adjustment to Kifer’s proof for the purpose of defining the pre-image topological pressure.

**Lemma 2.** For each \(x \in X_\epsilon\), \(k, n \in N\) with \(k \geq n\), and a nonrandom small positive number \(\epsilon\), the function \(P_{\text{pre}, n, \omega}(T, f, \epsilon, (T_k^{(\omega)})^{-1}x)\) is measurable in \(\omega\), and for any \(\delta > 0\), there exists a family of maximal \((\omega, n, \epsilon)\)-separated sets \(G_\omega \subset (T_k^{(\omega)})^{-1}x \subset E_\omega\) satisfying

\[
\sum_{y \in G_\omega} \exp S_n f(\omega, y) \geq (1 - \delta) P_{\text{pre}, n, \omega}(T, f, \epsilon, (T_k^{(\omega)})^{-1}x)
\]  

and depending measurably on \(\omega\) in the sense that \(G = \{ (\omega, x) : x \in G_\omega \} \in \mathcal{F} \times \mathcal{B}\), which also means that the mapping \(\omega \to G_\omega\) is measurable with respect to the Borel \(\sigma\)-algebra induced by the Hausdorff topology on the \(K(X)\) of compact subsets of \(X\). In particular, the supremum in the definition of \(P_{\text{pre}, n, \omega}(T, f, \epsilon, (T_k^{(\omega)})^{-1}x)\) can be taken only over families of \((\omega, n, \epsilon)\)-separated sets, which are measurable in \(\omega\).

**Proof.** Fix \(x \in X_\epsilon\). For \(q, n \in N_+\), let

- \(D_q = \{ (\omega, x_1, \ldots, x_q) : \omega \in \omega, x_i \in (T_k^{(\omega)})^{-1}x, \forall i \}\),
- \(E_q^n = \{ (\omega, x_1, \ldots, x_q) \in D_q : d_n^n(x_i, x_j) > \epsilon, \forall i \neq j \}\),
- \(E_q^n.l = \{ (\omega, x_1, \ldots, x_q) \in D_q : d_n^n(x_i, x_j) \geq \epsilon + 1/l, \forall i \neq j \}\),
- \(E_q^n(\omega) = \{ (x_1, \ldots, x_q) : (\omega, x_1, \ldots, x_q) \in E_q^n \}\)
- \(E_q^n.l(\omega) = \{ (x_1, \ldots, x_q) : (\omega, x_1, \ldots, x_q) \in E_q^n.l \}\).
Observe that $D_q \in \mathcal{F} \times \mathcal{B}^q$, where $\mathcal{B}^q$ is the product $\sigma$-algebra on the product of $q$ copies of $X$. This follows from [1, Theorem III.30] since

$$d_q((x_1, \ldots, x_q), (y_1, \ldots, y_q)) = \sum_{i=1}^{q} d(x_i, y_i)$$

is the distance function on $X^q$; and if $\mathcal{E}_\omega^{(q)}(x, k)$ denotes the product of $q$ copies of $\mathcal{E}_\omega(x, k) = \{y : (\omega, y) \in \mathcal{E}(x, k)\}$, then

$$d_q((x_1, \ldots, x_q), \mathcal{E}_\omega^{(q)}(x, k)) = \sum_{i=1}^{q} d(x_i, \mathcal{E}_\omega(x, k))$$

is measurable in $\omega$ for each $(x_1, \ldots, x_q) \in X^q$. Next we define $q(q-1)/2$ measurable functions $\psi_{ij}, 1 \leq i < j \leq q$, on $D_q$ by $\psi_{ij}(\omega, x_1, \ldots, x_q) = d_n(x_i, x_j)$. Then

$$E_{q,n,l} = \bigcap_{1 \leq i < j \leq q} \psi_{ij}^{-1}[\epsilon + 1/l, \infty) \in \mathcal{F} \times \mathcal{B}^q.$$ 

By the continuity of the RDS $T$, each $E_{q,n,l}(\omega)$ is a closed subset of $\mathcal{E}_\omega^{(q)}(x, k)$, thus it is compact. Clearly, $E_{q,n,l} \uparrow E_{q,n}^\omega$ and $E_{q,n,l}(\omega) \uparrow E_{q,n}^\omega(\omega)$ as $l \to \infty$. In particular, $E_{q,n}^\omega \in \mathcal{F} \times \mathcal{B}^q$.

Let $s_n(\omega, \epsilon)$ be the largest cardinality of all $(\omega, n, \epsilon)$-separated set in $\mathcal{E}_\omega(x, k)$ and $t_{n,l}(\omega, \epsilon) = \max\{q : E_{q,n,l}(\omega) \neq \emptyset\}$. Then by [1, Theorem III.23], we have

$$\{\omega : t_{n,l}(\omega, \epsilon) \geq q\} = \{\omega : E_{q,n,l}(\omega) \neq \emptyset\} = \text{Pr}_\Omega E_{q,n,l}^\omega \in \mathcal{F},$$

where $\text{Pr}_\Omega$ is the projection of $\Omega \times X^q$ to $\Omega$. Thus $t_{n,l}(\omega, \epsilon)$ is measurable in $\omega$. Now we get

$$\{\omega : s_n(\omega, \epsilon) \geq q\} = \{\omega : E_{q,n}(\omega) \neq \emptyset\} = \bigcup_{m=1}^{\infty} \bigcap_{l=m}^{\infty} \{\omega : E_{q,n,l}^\omega(\omega) \neq \emptyset\} \in \mathcal{F}.$$

Then $s_n(\omega, \epsilon)$ is measurable in $\omega$ as well. Since $s_n(\omega, \epsilon) \geq t_{n,l}(\omega, \epsilon)$ for all $l \geq 1$, then $t_{n,l}(\omega, \epsilon) \uparrow s_n(\omega, \epsilon)$ as $l \to \infty$, thus $t_{n,l}(\omega, \epsilon) = s_n(\omega, \epsilon)$ for all $l$ large enough (depending on $n$ and $\omega$). By [1, Lemma III.39], each function

$$g_{q,l} = \sup\{\sum_{i=1}^{q} \exp(S_n f(\omega, x_i)) : (x_1, \ldots, x_q) \in E_{q,n}^{n,l}(\omega)\}$$

8
is measurable. Thus the functions \( g_q = \sup_{i \geq 1} g_{q,i} \) and \( P_{\omega,n} (T, f, \epsilon, (T_k^k)^{-1} x) = \max_{1 \leq q \leq s(n, \epsilon)} g_q (\omega) \) are both measurable.

For each constant \( \delta > 0 \), the set

\[
F_{q,\delta}^n = \{ (\omega, x_1, \ldots, x_q) \in D_q : \sum_{i=1}^{q} \exp (S_n f (\omega, x_i)) \geq (1 - \delta) g_q (\omega) \}
\]

belongs to \( \mathcal{F} \times \mathcal{B}^q \). Hence

\[
G_{q,\delta}^n = F_{q,\delta}^n \cap E_q^n \in \mathcal{F} \times \mathcal{B}^q \quad \text{and} \quad G_{q,\delta}^{n,l} = F_{q,\delta}^n \cap E_{q}^{n,l} \in \mathcal{F} \times \mathcal{B}^q.
\]

Let

\[
G_{q,\delta}^n (\omega) = \{ (x_1, \ldots, x_q) : (\omega, x_1, \ldots, x_q) \in G_{q,\delta}^n \},
\]

\[
G_{q,\delta}^{n,l} (\omega) = \{ (x_1, \ldots, x_q) : (\omega, x_1, \ldots, x_q) \in G_{q,\delta}^{n,l} \}.
\]

Observe that \( G_{q,\delta}^{n,l}(\omega) \) are compact sets and \( G_{q,\delta}^{n,l}(\omega) \uparrow G_{q,\delta}^n(\omega) \) as \( l \to \infty \). The sets

\[
\tilde{\Omega}_{q,l} = \{ \omega : t_{n,l}(\omega, \epsilon) = s_n(\omega, \epsilon) = q \} \cap \{ \omega : G_{q,\delta}^{n,l}(\omega) \neq \emptyset \}
\]

are, clearly, measurable, and the sets \( \Omega_{q,l} = \tilde{\Omega}_{q,l} \setminus \tilde{\Omega}_{q,l-1}, l = 1, 2, \ldots, \) with \( \tilde{\Omega}_{q,0} = \emptyset \) are measurable, disjoint and \( \bigcup_{q,l \geq 1} \Omega_{q,l} = \Omega \). Thus (See [1, Theorem III.30]), the multifunction \( \Psi_{q,l,\delta} \) defined by \( \Psi_{q,l,\delta}(\omega) = G_{q,\delta}^{n,l}(\omega) \) for \( \omega \in \Omega_{q,l} \), is measurable, and it admits a measurable selection \( \sigma_{q,l,\delta} \) which is measurable map \( \sigma_{q,l,\delta} : \Omega_{q,l} \to X^q \) such that \( \sigma_{q,l,\delta}(\omega) \in G_{q,\delta}^{n,l}(\omega) \) for all \( \omega \in \Omega_{q,l} \). Let \( \zeta_q \) be the multifunction from \( X^q \) to \( q \)-point subsets of \( X \) defined by \( \zeta_q(1, \ldots, x_q) = \{ x_1, \ldots, x_q \} \subset X \). Then \( \zeta_q \circ \sigma_{q,l,\delta} \) is a multifunction assigning to each \( \omega \in \Omega_{q,l} \) a maximal \((\omega, n, \epsilon)\)-separated set \( G_{\omega} \subset (T_{\omega}^k)^{-1} x \) in \( E_{\omega} \) for which (2) holds true.

For any open set \( U \subset X \), let \( V_{U}^q(i) = \{ (x_1, \ldots, x_q) \in X^q : x_i \in U \} \), which is an open set of \( X^q \). Then, clearly,

\[
\{ \omega \in \Omega_{q,l} : \zeta_q \circ \sigma_{q,l,\delta}(\omega) \cap U \neq \emptyset \} = \bigcup_{i=1}^{q} \sigma_{q,l,\delta}^{-1} V_{U}^q(i) \in \mathcal{F}.
\]

Now we define the random variable \( m_q \) by \( m_q(\omega) = l \) for all \( \omega \in \Omega_{q,l} \). Let \( \Phi_{\tilde{s}} (\omega) = \zeta_{s_n(\omega, \epsilon)} \circ \sigma_{s_n(\omega, \epsilon), m_q(\omega), \delta}(\omega) \); then

\[
\{ \omega : \Phi_{\tilde{s}}(\omega) \cap U \neq \emptyset \} = \bigcup_{q,l=1}^{\infty} \{ \omega \in \Omega_{q,l} : \zeta_q \circ \sigma_{q,l,\delta}(\omega) \cap U \neq \emptyset \} \in \mathcal{F}.
\]
Hence $\Phi_\delta$ is a measurable multifunction which assigns to each $\omega \in \Omega$ a maximal $(\omega, n, \epsilon)$-separated set $G_{\omega}$ for which (2) holds and Lemma 2 follows since $\delta > 0$ is arbitrary.

In view of this assertion, for each $f \in L^1_\mathcal{E}(\Omega, \mathcal{C}(X))$ and any positive number $\epsilon$ we can introduce the function

$$P_{\text{pre}}(T, f, \epsilon) = \limsup_{n \to \infty} \frac{1}{n} \sup_{k \geq n} \sup_{x \in X_{\mathcal{E}}} \int \log P_{\text{pre}, n, \omega}(T, f, \epsilon, (T_{\omega}^k)^{-1}x) \, dP(\omega). \quad (3)$$

Note that though we set $x \in X_{\mathcal{E}}$, only those points such that $x \in \mathcal{E}_{T^k \omega}$ act in the function $P_{\text{pre}}(T, f, \epsilon)$.

**Definition 3.** The pre-image topological pressure of a function $f \in L^1_\mathcal{E}(\Omega, \mathcal{C}(X))$ for bundle $\text{RDS } T$ is the map $P_{\text{pre}}(T, \cdot) : L^1_\mathcal{E}(\Omega, \mathcal{C}(X)) \to \mathbb{R} \cup \{\infty\}$, where $P_{\text{pre}}(T, f) = \lim_{\epsilon \to 0} P_{\text{pre}}(T, f, \epsilon)$. The limit exists since $P_{\text{pre}}(T, f, \epsilon)$ is monotone in $\epsilon$ and, in fact, $\lim_{\epsilon \to 0}$ above equals $\sup_{\epsilon > 0}$.

**Definition 4.** The pre-image topological entropy for bundle $\text{RDS } T$ is defined as

$$h^{(r)}_{\text{pre}}(T) = P_{\text{pre}}(T, 0) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \sup_{k \geq n} \sup_{x \in X_{\mathcal{E}}} \int \log s_n(\omega, \epsilon, (T_{\omega}^k)^{-1}x) \, dP(\omega),$$

where $s_n(\omega, \epsilon, (T_{\omega}^k)^{-1}x)$ is the largest cardinality of an $(\omega, n, \epsilon)$-separated set of $(T_{\omega}^k)^{-1}x$.

**Remark 5.** Definition 4 is different from the random pre-image topological entropy defined by Zhu [27], which in our terminology can be expressed as

$$h^{(r)}_{\text{pre}}(T) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \sup_{x \in X_{\mathcal{E}}} \int \log \sup_{\epsilon > 0} s_n(\omega, \epsilon, (T_{\omega}^k)^{-1}x) \, dP(\omega).$$

The measurability of the function $\sup_{x \in X_{\mathcal{E}}} s_n(\omega, \epsilon, (T_{\omega}^k)^{-1}x)$ can not be guaranteed in most cases. In Definition 4, based on Lemma 2, $s_n(\omega, \epsilon, (T_{\omega}^k)^{-1}x)$ is always measurable. On the other hand, through a rigorous investigation, it is not hard to see that if we replace Zhu’s definition by Definition 4 and make a little change to Zhu’s argument, then the variational principle for the pre-image entropy for bundle RDS $T$ still holds, namely, $h^{(r)}_{\text{pre}}(T) = \sup \{ h^{(r)}_{\text{pre}, \mu}(T) : \mu \in \mathcal{M}_1(\mathcal{E}, T) \}$. In fact, we just need to make some adjustment to the order of the supremum and the logarithm and restrict the whole space $\Omega \times X$ to the measurable subset $\mathcal{E}$ in his proof.
Remark 6. If the measure space $\Omega$ consists of a single point, i.e., $\Omega = \{\omega\}$, then bundle RDS $T$ reduces to a deterministic dynamical system $(\mathcal{E}_\omega, d, T)$, where $T : \mathcal{E}_\omega \to \mathcal{E}_\omega$ is continuous. Furthermore, if $\mathcal{E}_\omega = X$, then Definition 3 is just the pre-image topological pressure defined by Zeng [7] for deterministic dynamical systems except for the order of the supremum and the logarithm, and the difference between the two kinds of order does not affect the variational principle and the method of the argument.

For a given $m \in \mathbb{N}_+$, if we replace $\vartheta$ by $\vartheta^m$ and consider bundle RDS $T^m$ defined by $(T^m)^n = T_{\vartheta^{(n-1)m}_\omega}^m \cdots T_{\vartheta^m_\omega}^m T^m_\omega$, i.e., $(T^m)^n = T^m_\omega$, then the pre-image topological pressure has the following power rule.

**Proposition 7.** For any $m > 0$, $P_{pre}(T^m, S_m f) = m P_{pre}(T, f)$.

**Proof.** Fix $m \in \mathbb{N}$. Let $n \in \mathbb{N}$, $k \geq n$ and $x \in X_\vartheta$. If $E$ is an $(\omega, n, \epsilon)$-separated set of $(T^m)^k_\omega x$ for $T^m$, then $E$ is also an $(\omega, mn, \epsilon)$-separated set of $(T^{mk})^{-1}x$ for $T$. Since $(T^m)^n = T^m_\omega$, so

$$P_{pre, n, \omega}(T^m, S_m f, \epsilon, ((T^m)^k_\omega)^{-1}x) \leq P_{pre, mn, \omega}(T, f, \epsilon, (T^{mk})^{-1}x)$$

Hence

$$P_{pre}(T^m, S_m f, \epsilon) = \limsup_{n \to \infty} \frac{1}{n} \sup_{k \geq n} \sup_{x \in X_\vartheta} \int \log P_{pre, n, \omega}(T^m, S_m f, \epsilon, ((T^m)^k_\omega)^{-1}x) d\mathcal{P}(\omega)$$

$$\leq \limsup_{n \to \infty} \frac{1}{n} \sup_{k \geq mn} \sup_{x \in X_\vartheta} \int \log P_{pre, mn, \omega}(T, f, \epsilon, (T^{mk})^{-1}x) d\mathcal{P}(\omega)$$

$$\leq mn \limsup_{n \to \infty} \frac{1}{mn} \sup_{k \geq mn} \sup_{x \in X_\vartheta} \int \log P_{pre, mn, \omega}(T, f, \epsilon, (T^{mk})^{-1}x) d\mathcal{P}(\omega)$$

$$\leq \limsup_{n \to \infty} \frac{1}{n} \sup_{k \geq n} \sup_{x \in X_\vartheta} \int \log P_{pre, n, \omega}(T, f, \epsilon, (T^{mk})^{-1}x) d\mathcal{P}(\omega)$$

$$\leq m P_{pre}(T, f, \epsilon).$$

Therefore, $P_{pre}(T^m, S_m f) \leq m P_{pre}(T, f)$.

For any $\epsilon > 0$, by the continuity of $T^m$, there exists some small enough $\delta > 0$ such that if $d(y, z) \leq \delta$, $y, z \in \mathcal{E}_\omega$, then $d^m_\omega(y, z) \leq \epsilon$. For any positive integer $n$, there exists some integer $l$ such that $mn \leq l < m(n + 1)$. It is easy to see that any $(\omega, l, \epsilon)$-separated set of $(T^k_\omega)^{-1}x$ for $T$ is also an
(ω, n, δ)-separated set of $(T^k_ω)^{-1}x$ for $T^m$. For $k \geq l$, let $k = k_1m + q$, where $k_1, q \in \mathbb{N}, 0 \leq q < m$, then $k_1 \geq n$. Let $x' = T_{δ(k+m-q-1)}\cdots T_{δk}x$, then $x \in (T_{δ(k+m-q-1)}\cdots T_{δk})^{-1}x'$, and then $(T^k_ω)^{-1}x = (T^m_ω)^{-1}x \subset (T^m_ω)^{(k+1)}^{-1}x' = ((T^m_ω)^{k'})^{-1}x'$, where $k' = k_1 + 1$. Hence, any $(ω, n, δ)$-separated set of $(T^k_ω)^{-1}x$ for $T^m$ is also an $(ω, n, δ)$-separated set of $((T^m_ω)^{k'})^{-1}x'$ for $T^m$. Therefore, for $k \geq l$,

$$P_{pre,l,ω}(T, f, ϵ, (T^k_ω)^{-1}x)$$

$$= \sup \{ \sum \exp S_l f(ω, y) : E \text{ is an } (ω, l, ϵ) \text{-separated set of } (T^k_ω)^{-1}x \text{ for } T \}$$

$$= \sup \{ \sum \exp (\sum_{i=0}^{n-1} S_m f(ω, y, T^m y) + \sum_{j=mn}^{l-1} f(ω, y, T^m y)) : E \text{ is an } (ω, l, ϵ) \text{-separated set of } (T^k_ω)^{-1}x \text{ for } T \}$$

$$\leq \exp \sum_{j=mn}^{l-1} \| f(ω, y) \|_∞ \sup \{ \sum \exp (\sum_{i=0}^{n-1} S_m f(ω, y, T^m y) : E \text{ is an } (ω, n, δ) \text{-separated set of } (T^k_ω)^{-1}x \text{ for } T \}$$

$$\leq \exp \sum_{j=mn}^{l-1} \| f(ω, y) \|_∞ \sup \{ \sum \exp (\sum_{i=0}^{n-1} S_m f(ω, y, T^m y) : E \text{ is an } (ω, n, δ) \text{-separated set of } ((T^m_ω)^{k'})^{-1}x' \text{ for } T^m \}$$

$$= \exp \sum_{j=mn}^{l-1} \| f(ω, y) \|_∞ P_{pre,n,ω}(T^m, S_m f, δ, ((T^m_ω)^{k'})^{-1}x').$$

Since $f \in \mathbf{L}_1^δ(Ω, C(X))$, so $\int \sum_{j=mn}^{l-1} \| f(ω, y) \|_∞ dP(ω) ≤ m\|f\| < ∞$. Then by the definition of $P_{pre}(T, f, ϵ)$, we have

$$P_{pre}(T, f, ϵ)$$

$$= \lim_{l → ∞} \frac{1}{l} \sup \sum \sup \int_0 \log P_{pre,l,ω}(T, f, ϵ, (T^k_ω)^{-1}x) dP(ω)$$

$$\leq \lim_{l → ∞} \frac{1}{mn} \sup \sum \sup \int_0 \left( \sum_{j=mn}^{l-1} \| f(ω, y) \|_∞ + \log P_{pre,n,ω}(T^m, S_m f, δ, ((T^m_ω)^{k'})^{-1}x')) dP(ω)$$

$$= \frac{1}{m} \lim_{l → ∞} \frac{1}{n} \sup \sum \sup \int_0 \log P_{pre,n,ω}(T^m, S_m f, δ, ((T^m_ω)^{k'})^{-1}x') dP(ω)$$
= \frac{1}{m} P_{\text{pre}}(T^m, S_m f, \delta).

Hence,

\[ m P_{\text{pre}}(T, f, \epsilon) \leq P_{\text{pre}}(T^m, S_m f, \delta). \]

If \( \epsilon \to 0 \), then \( \delta \to 0 \). Hence we obtain \( m P_{\text{pre}}(T, f) \leq P_{\text{pre}}(T^m, S_m f) \) and complete the proof.

\[ \Box \]

4. Pre-image variational principle for bundle RDS

**Theorem 8.** If \( T \) is a continuous bundle RDS on \( \mathcal{E} \) and \( f \in L^1_\mathcal{E}(\Omega, \mathcal{C}(X)) \), then

\[ P_{\text{pre}}(T, f) = \sup \{ h_{\text{pre}, \mu}(T) + \int f \, d\mu : \mu \in \mathcal{M}_1^1(\mathcal{E}, T) \}. \]

**Proof.** (1) Let \( \mu \in \mathcal{M}_1^1(\mathcal{E}, T) \), \( \mathcal{P} = \{ P_1, \cdots, P_k \} \) be a finite measurable partition of \( X \), and \( \epsilon \) be a positive number with \( \epsilon k \log k < 1 \). Let \( \mathcal{P}(\omega) = \{ P_1(\omega), \cdots, P_k(\omega) \} \) be the corresponding partition of \( \mathcal{E}_\omega \), where \( P_i(\omega) = P_i \cap \mathcal{E}_\omega, i = 1, \cdots, k \). By the regularity of \( \mu \), we can find compact sets \( Q_i \subseteq P_i, 1 \leq i \leq k \), such that

\[ \int \mu_\omega(P_i(\omega) \setminus Q_i(\omega)) \, d\mathcal{P}(\omega) < \epsilon, \]

where \( Q_i(\omega) = Q_i \cap \mathcal{E}_\omega \). Let \( \mathcal{Q}(\omega) = \{ Q_0(\omega), \cdots, Q_k(\omega) \} \) be the partition of \( \mathcal{E}_\omega \), where \( Q_0(\omega) = \mathcal{E}_\omega \setminus \bigcup_{i=1}^k Q_i(\omega) \). Then by the proof of [27, Theorem 4] and [24, Lemma II. 1.3], the following inequality holds (See [27] for details),

\[ h_{\text{pre}, \mu}^{(r)}(T, \Omega \times \mathcal{P}) \leq h_{\text{pre}, \mu}^{(r)}(T, \Omega \times \mathcal{Q}) + 1, \tag{4} \]

where \( \Omega \times \mathcal{P} \) (respectively by \( \Omega \times \mathcal{Q} \)) is the partition of \( \Omega \times X \) into sets \( \Omega \times P_i \) (respectively by \( \Omega \times Q_i \)).

Let

\[ \mathcal{Q}_n(\omega) = \bigvee_{i=0}^{n-1} (T^i_\omega)^{-1} \mathcal{Q}(\theta^i \omega). \]

For each \( k \), we choose a non-decreasing sequence of finite partitions \( \beta_1^k \leq \beta_2^k \leq \cdots \) with diameters tending to zero for which \( \mathcal{B}_{\theta^k \omega} = \bigcup_{j=1}^\infty \beta_j^k \) up to set of measure 0, and satisfying

\[ H_{\mu_\omega}(\mathcal{Q}_n(\omega) \setminus (T^k_\omega)^{-1} \mathcal{B}_{\theta^k \omega}) = \lim_{j \to \infty} H_{\mu_\omega}(\mathcal{Q}_n(\omega) \setminus (T^k_\omega)^{-1} \beta_j^k). \]
Let $\eta = \{Q_0 \cup Q_1, \ldots, Q_0 \cup Q_k\}$, and $\tau$ be the Lebesgue number of $\eta$. We pick a small nonrandom $\delta$ with $0 < 4\delta < \tau$. Let $\epsilon_1 := \epsilon_1(\omega, n, \epsilon) > 0$ for each $\omega \in \Omega$, such that if $d(x, y) < \epsilon_1$, $x, y \in E_\omega$, then $d_n^\omega(x, y) < \delta$.

For each $\omega \in \Omega$, $\{(T^k_\omega)^{-1} x : x \in E_{\vartheta^k \omega}\}$ is an upper semi-continuous decomposition of $E_\omega$. Thus for each $x \in E_{\vartheta^k \omega}$, there is an $\epsilon_2 := \epsilon_2(\omega, x, k, \epsilon_1)$ such that if $d(x, y) < \epsilon_2$, $y \in E_{\vartheta^k \omega}$ and $y_1 \in (T^k_\omega)^{-1} y$, then there is an $x_1 \in (T^k_\omega)^{-1} x$ such that $d(x_1, y_1) < \epsilon_1$. Let $U_k$ be the collection of open $\epsilon_2$ balls in $E_{\vartheta^k \omega}$ as $x$ varies in $E_{\vartheta^k \omega}$ and $\epsilon_3$ be the Lebesgue number for $U_k$.

Since $\text{diam}(\beta^k_j) \to 0$ as $j \to \infty$, we can choose large enough $j$ such that for each $B \in \beta^k_j$, $\text{diam} B < \epsilon_3$. For each $A \in (T^k_\omega)^{-1} \beta^k_j$, let $\mu_{\omega, A}$ be the conditional measure of $\mu_{\omega}$ restricted to $A$. $Q_n(\omega, A) = \{A \cap C : C \in Q_n(\omega), A \cap C \neq \emptyset\}$, $S_n^f(\omega, A \cap C) = \sup \{S_n f(\omega, x) : x \in A \cap C\}$. Then by the standard inequality

$$\sum_{i=1}^m p_i(a_i - \log p_i) \leq \log \sum_{i=1}^m \exp a_i,$$

for any probability vector $(p_1, \ldots, p_m)$, we have

$$H_{\mu_\omega}(Q_n(\omega)|{(T^k_\omega)^{-1} \beta^k_j}) + \int_{E_\omega} S_n f \, d\mu_\omega$$

$$= \sum_{A \in (T^k_\omega)^{-1} \beta^k_j} (\mu_\omega(A) H_{\mu_{\omega, A}}(Q_n(\omega, A)) + \int_A S_n f \, d\mu_\omega)$$

$$\leq \sum_{A \in (T^k_\omega)^{-1} \beta^k_j} \mu_\omega(A) \sum_{A \cap C \in Q_n(\omega, A)} \mu_{\omega, A}(A \cap C) \left(-\log \mu_{\omega, A}(A \cap C) + S_n^f(\omega, A \cap C)\right)$$

$$\leq \max_{A \in (T^k_\omega)^{-1} \beta^k_j} \log \sum_{A \cap C \in Q_n(\omega, A)} \exp S_n^f(\omega, A \cap C).$$

(5)

In the sequel, let $A$ be the maximal one satisfying the above inequality and $B \in \beta^k_j$ with $A = (T^k_\omega)^{-1} B$. For each $A \cap C \in Q_n(\omega, A)$, we choose some point $x_C \in A \cap C$ such that $S_n f(\omega, x_C) = S_n^f(\omega, A \cap C)$.

Since $T^k_\omega(x_C) \in \overline{B}$, and $\text{diam} B < \epsilon_3$, there is an $u_B \in E_{\vartheta^k \omega}$ such that $d(u_B, T^k_\omega(x_C)) < \epsilon_2$. Hence there is $y_1 \in (T^k_\omega)^{-1} u_B$ such that $d(x_C, y_1) < \epsilon_1$ and then $d_n^\omega(x_C, y_1) < \delta$.

Let $E^\omega_A$ be a maximal $(\omega, n, \delta)$-separated set in $(T^k_\omega)^{-1} u_B$. Since $E^\omega_A$ is also a spanning set, there is a point $z(x_C) \in E^\omega_A$ such that $d_n^\omega(y_1, z(x_C)) \leq \delta$, 

14
then $d_n^ω(x_C, z(x_C)) \leq 2\delta$ and hence

\begin{equation}
S_n^ωf(ω, A \cap C) \leq S_n f(ω, z(x_C)) + \sum_{i=0}^{n-1} \kappa_{2\delta}^{(j)}(ψ^i ω).
\end{equation}

(6)

We now show that if $y \in E_ω^A$, then

\text{card}\{A \cap C \in Q_n(ω, A) : z(x_C) = y\} \leq 2^n.

(7)

Suppose $z(x_C) = z(x_C)$. There exist $x_C \in A \cap C$ and $x_C \in A \cap C$ so that $d_n^ω(x_C, z(ω)) < 2\delta$ and $d_n^ω(x_C, z(ω)) < 2\delta$. Therefore $d_n^ω(x_C, x_C) < 4\delta$; so $T_n^ω(x_C)$ and $T_n^ω(x_C)$ are in the same element of $η$, say $Q_0 \cup Q_j, 0 \leq i < n$. Hence there are at most $2^n$ elements $A \cap C$ of $Q_n(ω, A)$ equal to a fixed member of $E_ω^A$.

Combining (6) and (7), we get

$$\sum_{A \cap C \in Q_n(ω, A)} \exp S_n^ωf(ω, A \cap C) \leq 2^n \sum_{y \in E_ω^A} \exp (S_n f(ω, y) + \sum_{i=0}^{n-1} \kappa_{2\delta}^{(j)}(ψ^i ω)).$$

Taking the logarithm of both parts and using the resulting inequality in order to estimate the righthand side of (5), we obtain

$$H_{µ_ω}(Q_n(ω)|(T_n^ω)^{-1} E_ω) + \int \limits_{E_ω} S_n f dµ_ω$$

$$\leq n \log 2 + \log \sum_{y \in E_ω^A} \exp S_n f(ω, y) + \sum_{i=0}^{n-1} \kappa_{2\delta}^{(j)}(ψ^i ω)$$

$$\leq n \log 2 + \log P_{pre,n,ω}(T, f, \delta, (T_n^ω)^{-1} u_B) + \sum_{i=0}^{n-1} \kappa_{2\delta}^{(j)}(ψ^i ω).$$

Integrating this with respect to $P$ and letting $j \to \infty$, we have

$$\int H_{µ_ω}(Q_n(ω)|(T_n^ω)^{-1} E_ω) dP(ω) + \int S_n f dµ$$

$$\leq n \log 2 + \int \log P_{pre,n,ω}(T, f, \delta, (T_n^ω)^{-1} u_B) dP(ω) + \int \sum_{i=0}^{n-1} \kappa_{2\delta}^{(j)}(ψ^i ω) dP(ω)$$

$$\leq n \log 2 + \sup_{x \in E_ρ_ω} \int \log P_{pre,n,ω}(T, f, \delta, (T_n^ω)^{-1} x) dP(ω) + \int \sum_{i=0}^{n-1} \kappa_{2\delta}^{(j)}(ψ^i ω) dP(ω).$$
Letting $k \to \infty$, dividing by $n$ and letting $n \to \infty$, by Proposition 1 and the equality (3), we have
\begin{equation}
 h^{(r)}_{\text{pre, } \mu}(T, \Omega \times \mathcal{Q}) + \int f \, d\mu \leq \log 2 + P_{\text{pre}}(T, f, \delta) + \int \kappa^{(f)}_{23} \, d\mathbf{P}(\omega).
\end{equation}

Using (4), we derive the inequality
\begin{equation}
 h^{(r)}_{\text{pre, } \mu}(T, \Omega \times \mathcal{P}) + \int f \, d\mu \leq \log 2 + 1 + P_{\text{pre}}(T, f, \delta) + \int \kappa^{(f)}_{23} \, d\mathbf{P}(\omega).
\end{equation}

Since this is true for all finite partitions $\mathcal{P}$ of $X$ and all positive $\delta$, we have
\begin{equation}
 h^{(r)}_{\text{pre, } \mu}(T) + \int f \, d\mu \leq \log 2 + 1 + P_{\text{pre}}(T, f).
\end{equation}

Since $h^{(r)}_{\text{pre, } \mu}(T^n) = mh^{(r)}_{\text{pre, } \mu}(T)$ (See [27, Proposition 4]), the same arguments as above applied to $T^n$ and to $S_nf$ yield
\begin{equation}
 n(h^{(r)}_{\text{pre, } \mu}(T) + \int f \, d\mu) \leq \log 2 + 1 + P_{\text{pre}}(T^n, S_nf).
\end{equation}

Taking into account Proposition 7, dividing by $n$ and letting $n \to \infty$, we conclude that
\begin{equation}
 h^{(r)}_{\text{pre, } \mu}(T) + \int f \, d\mu \leq P_{\text{pre}}(T, f).
\end{equation}

(2) By the equality 3, we can choose a sequence $n_i \to \infty$, $k_i \geq n_i$, and points $x_{\varphi^{k_i}\omega} \in \mathcal{E}_{\varphi^{k_i}\omega}$ for each $\omega \in \Omega$ such that
\begin{equation}
 P_{\text{pre}}(T, f, \epsilon) = \lim_{i \to \infty} \frac{1}{n_i} \int \log P_{\text{pre, } n_i, \omega}(T, f, \epsilon, (T_{\omega}^{k_i})^{-1} x_{\varphi^{k_i}\omega}) \, d\mathbf{P}(\omega).
\end{equation}

For a small nonrandom $\epsilon > 0$, by Lemma 2, we can choose a family of maximal $(\omega, n_i, \epsilon)$-separated sets $G(\omega, n_i, \epsilon) \subset (T_{\omega}^{k_i})^{-1} x_{\varphi^{k_i}\omega} \subset \mathcal{E}_{\omega}$, which are measurable in $\omega$, such that
\begin{equation}
 \sum_{x \in G(\omega, n_i, \epsilon)} \exp S_{n_i} f(\omega, x) \geq \frac{1}{\epsilon} P_{\text{pre, } n_i, \omega}(T, f, \epsilon, (T_{\omega}^{k_i})^{-1} x_{\varphi^{k_i}\omega}).
\end{equation}
Next, we define probability measures $\nu^{(i)}$ on $\mathcal{E}$ via their measurable disintegrations

$$\nu^{(i)}_\omega = \frac{\sum_{x \in G(\omega, n_i, \epsilon)} \exp S_{n_i} f(\omega, x) \delta_x}{\sum_{x \in G(\omega, n_i, \epsilon)} \exp S_{n_i} f(\omega, x)}$$

so that $d\nu^{(i)}(\omega, x) = d\nu^{(i)}_\omega(x) d\mathbf{P}(\omega)$.

Let

$$\mu^{(i)} = \frac{1}{n_i} \sum_{j=0}^{n_i-1} \Theta^j \nu^{(i)}.$$

Then, by Lemma 2.1 (i)-(ii) of [26], we can choose a subsequence $n_{i_1}$ of $\{n_i\}$ such that $\lim_{i \to \infty} \mu^{(i_1)} = \mu$ for some $\mu \in \mathcal{M}_T^1(\mathcal{E}, T)$. Without loss of generality, we still assume that $\lim_{i \to \infty} \mu^{(i)} = \mu$.

Next, we choose a partition $\mathcal{P} = \{P_1, \cdots, P_k\}$ of $X$ with $\text{diam} \mathcal{P} \leq \epsilon$, which satisfies $\int \mu(\partial P_i) \, d\mathbf{P}(\omega) = 0$ for all $1 \leq i \leq k$, where $\partial$ denotes the boundary. Let $\mathcal{P}(\omega) = \{P_1(\omega), \cdots, P_k(\omega)\}$, $P_i(\omega) = P_i \cap \mathcal{E}_\omega$. Since each element of $\bigvee_{l=0}^{n_i-1}(T^l_\omega)^{-1} \mathcal{P}(\vartheta^l \omega)$ contains at most one element of $G(\omega, n_i, \epsilon)$, then by the inequality (8), we have

$$H_{\nu^{(i)}_\omega} \left( \bigvee_{l=0}^{n_i-1}(T^l_\omega)^{-1} \mathcal{P}(\vartheta^l \omega) \right) |_{(T^l_\omega)^{-1} x_{\vartheta^l \omega}} + \int S_{n_i} f(\omega, x) \, d\nu^{(i)}_\omega$$

$$\geq \log P_{\text{pre}, n_i, \omega}(T, f, \epsilon, (T^l_\omega)^{-1} x_{\vartheta^l \omega}) - 1.\quad(9)$$

For $\omega \in \Omega$, let $\mathcal{C}_\omega$ be the subcollection of $\mathcal{B}_\omega^- \subset \mathcal{C}_\omega$ consisting of $\mu_\omega$-null sets. For any $\sigma$-algebra $\mathcal{A}$ of subsets of $\mathcal{E}_\omega$, there is an enlarged $\sigma$-algebra $\mathcal{A}_{\mathcal{C}_\omega}$ defined by $A \in \mathcal{A}_{\mathcal{C}_\omega}$ if and only if there are sets $B, M, N$ such that $A = B \cup M, B \in \mathcal{A}, N \in \mathcal{C}_\omega$ and $M \subset N$. The $\sigma$-algebra $\mathcal{B}_\omega^-$ is simply the standard $\mu_\omega$-completion of $\mathcal{B}_\omega^-$. Let $\mathcal{B}_\omega^k = ((T^k_\omega)^{-1} \mathcal{B}_{\vartheta^k \omega})_{\mathcal{C}_\omega}$ for all $k \geq 1$. Since $T_{\omega-1}^l \mathcal{C}_{\vartheta^l \omega} \subset \mathcal{C}_\omega$ for each $\omega$, we have that $\mathcal{B}_\omega^1 \supset \mathcal{B}_\omega^2 \supset \cdots$. Let $\mathcal{B}_\omega^\infty = \bigcap_{k \geq 1} \mathcal{B}_\omega^k$, $\mathcal{B}_\omega^- \subset \mathcal{B}_\omega^- \subset \mathcal{B}_\omega^\infty$ and $(T^l_\omega)^{-1} \mathcal{B}_{\vartheta^l \omega}^k \subset \mathcal{B}_\omega^{l+k}$ for all $l \geq 1$.

Similarly, let $\mathcal{C}_\mathcal{E}$ be the subcollection of $(\mathcal{F} \times \mathcal{B})_{\mathcal{C}_\mathcal{E}}^\infty$ consisting of $\mu$-null sets, $(\mathcal{F} \times \mathcal{B})_{\mathcal{C}_\mathcal{E}}^k = (\Theta^{-k}(\mathcal{F} \times \mathcal{B})_{\mathcal{C}_\mathcal{E}})_{\mathcal{C}_\mathcal{E}}$ and $(\mathcal{F} \times \mathcal{B})_{\mathcal{C}_\mathcal{E}}^\infty = \bigcap_{k \geq 1} (\mathcal{F} \times \mathcal{B})_{\mathcal{C}_\mathcal{E}}^k$. Clearly, $(\mathcal{F} \times \mathcal{B})_{\mathcal{C}_\mathcal{E}}^k \supset (\mathcal{F} \times \mathcal{B})_{\mathcal{C}_\mathcal{E}}^l \supset \cdots, (\mathcal{F} \times \mathcal{B})_{\mathcal{C}_\mathcal{E}}^k \subset (\mathcal{F} \times \mathcal{B})_{\mathcal{C}_\mathcal{E}}^l \subset (\mathcal{F} \times \mathcal{B})_{\mathcal{C}_\mathcal{E}}^\infty$ and
\( \Theta^{-l}(\mathcal{F}_E \vee (\mathcal{F} \times \mathcal{B})^k_E) \subset \mathcal{F}_E \vee (\mathcal{F} \times \mathcal{B})^{l+k} \) for all \( l \geq 1 \). Similar to the proof of Proposition 1, we conclude that for each \( k \) and any finite partition \( O = \{O_i\} \) of \( E \),

\[
H_{\nu^{(i)}}(O|\mathcal{F}_E \vee (\mathcal{F} \times \mathcal{B})^k_E) = \int H_{\nu^{(i)}}(O(\omega)|\mathcal{B}_E^k) \, dP(\omega), \tag{10}
\]

where \( O(\omega) = \{O_i(\omega)\} \), \( O_i(\omega) = \{x \in \mathcal{E}_\omega : (\omega, x) \in O_i\} \).

Note that \( \nu^{(i)}(O_i) \) is supported on \((T_{\omega}^{k_i})^{-1}x_{\varphi^{k_i}\omega}\), the canonical system of conditional measures induced by \( \nu^{(i)}(O_i) \) on the measurable partition \( \{(T_{\omega}^{k_i})^{-1}x|x \in \mathcal{E}_{\varphi^{k_i}\omega}\} \) reduces to a single measure on the set \((T_{\omega}^{k_i})^{-1}x_{\varphi^{k_i}\omega}\), which may be identified with \( \nu^{(i)}(O_i) \). Now, each element \( A \in \mathcal{B}_{\omega}^k \) can be expressed as the disjoint union \( A = B \cup C \) with \( B \in (T_{\omega}^{k_i})^{-1}\mathcal{B}_{\varphi^{k_i}\omega} \) and \( C \in \mathcal{C}_\omega \). Since \( \nu^{(i)}(O_i) \) is supported on elements of \((T_{\omega}^{k_i})^{-1}x_{\varphi^{k_i}\omega}\), we have \( \nu^{(i)}(O_i)(C) = 0 \). Hence for any finite partition \( \gamma \) of \( \mathcal{E}_\omega \), we have

\[
H_{\nu^{(i)}}(\gamma|\mathcal{B}_{\omega}^k) = H_{\nu^{(i)}}(\gamma|(T_{\omega}^{k_i})^{-1}x_{\varphi^{k_i}\omega}). \tag{11}
\]

Let \( Q = \{Q_1, \cdots, Q_k\} \), where \( Q_i = (\Omega \times P_i) \cap \mathcal{E} \), then \( Q \) is a partition of \( \mathcal{E} \) and \( Q_i(\omega) = \{x \in \mathcal{E}_\omega : (\omega, x) \in Q_i\} = P_i(\omega) \). Integrating in (9) with respect to \( P \), then by (10), (11) and \( \int S_n, f \, d\nu^{(i)} = n_i \int f \, d\mu^{(i)} \), we obtain the inequality

\[
\begin{align*}
H_{\nu^{(i)}} \left( \bigvee_{l=0}^{n_i-1} (\Theta^{l})^{-1}Q|\mathcal{F}_E \vee (\mathcal{F} \times \mathcal{B})^k_E \right) + n_i \int f \, d\mu^{(i)} \\
\geq \int \log P_{\text{pre},n_i,\omega}(T, f, \epsilon, (T_{\omega}^{k_i})^{-1}x_{\varphi^{k_i}\omega}) \, dP(\omega) - 1.
\end{align*}
\tag{12}
\]

For \( q, n_i \in \mathbb{N} \) with \( 1 < q < n_i \), let \( a(s) \) denote the integer part of \((n_i - s)q^{-1}\) for all \( 0 \leq s < q \). Then, clearly, for each \( s \), we have

\[
\bigvee_{l=0}^{n_i-1} (\Theta^{l})^{-1}Q = \bigvee_{r=0}^{a(s)-1} (\Theta^{r+q})^{-1} \bigvee_{l=0}^{q-1} (\Theta^{l})^{-1}Q \vee \bigvee_{l \in S} (\Theta^{l})^{-1}Q,
\]

where \( \text{card}S \leq 2q \).

Since \( \text{card}Q = k \), by the subadditivity of conditional entropy (See [24, Section 2.1]), we have

\[
H_{\nu^{(i)}} \left( \bigvee_{l=0}^{n_i-1} (\Theta^{l})^{-1}Q|\mathcal{F}_E \vee (\mathcal{F} \times \mathcal{B})^k_E \right)
\]

18
\[
\begin{align*}
&\leq \sum_{r=0}^{a(s)-1} H_{\mu(i)} \left( (\Theta^{rq+s})^{-1} \bigvee_{t=0}^{q-1} (\Theta^t)^{-1} Q|F_\mathcal{E} \vee (F \times B)^{k_i}_\mathcal{E} \right) + 2q \log k \\
&\leq \sum_{r=0}^{a(s)-1} H_{\mu(i)} \left( (\Theta^{rq+s})^{-1} \bigvee_{t=0}^{q-1} (\Theta^t)^{-1} Q|((\Theta^{rq+s})^{-1}(F \times B)^{k_i}_\mathcal{E}) \right) + 2q \log k \\
&= \sum_{r=0}^{a(s)-1} H_{\Theta^{rq+s}\mu(i)} \left( \bigvee_{t=0}^{q-1} (\Theta^t)^{-1} Q|F_\mathcal{E} \vee (F \times B)^{k_i}_\mathcal{E} \right) + 2q \log k.
\end{align*}
\]

Summing this inequality over \( s \in \{0, 1, \cdots, q - 1\} \), we get

\[
qH_{\mu(i)} \left( \bigvee_{t=0}^{q-1} (\Theta^t)^{-1} Q|F_\mathcal{E} \vee (F \times B)^{k_i}_\mathcal{E} \right)
\]

\[
\leq \sum_{t=0}^{n_i-1} H_{\Theta^{t}\mu(i)} \left( \bigvee_{t=0}^{q-1} (\Theta^t)^{-1} Q|F_\mathcal{E} \vee (F \times B)^{k_i}_\mathcal{E} \right) + 2q^2 \log k
\]

\[
\leq n_i H_{\mu(i)} \left( \bigvee_{t=0}^{q-1} (\Theta^t)^{-1} Q|F_\mathcal{E} \vee (F \times B)^{k_i}_\mathcal{E} \right) + 2q^2 \log k
\]

\[
\leq n_i H_{\mu(i)} \left( \bigvee_{t=0}^{q-1} (\Theta^t)^{-1} Q|F_\mathcal{E} \vee (F \times B)^{\infty}_\mathcal{E} \right) + 2q^2 \log k,
\]

where the second inequality, as in Kifer’s works [26], relies on the general property of conditional entropy of partition \( \sum_i p_i H_{\eta_i}(\xi|A) \leq H_{\sum_i p_i \eta_i}(\xi|A) \) which holds for any finite partition \( \xi \), \( \sigma \)-algebra \( A \), probability measures \( \eta_i \), and probability vector \( (p_i) \), \( i = 1, \ldots, n \), in view of the convexity of \( t \log t \) in the same way as that in the unconditional case (cf. [12, pp.183 and 188 ] ). This together with (12) yields

\[
q \int \log P_{\text{pre}, n_i, \omega}(T, f, \epsilon, (T^k_\omega)^{-1}x_{\eta^{n_i}_\omega}) \, dP(\omega) - q
\]

\[
\leq n_i H_{\mu(i)} \left( \bigvee_{t=0}^{q-1} (\Theta^t)^{-1} Q|F_\mathcal{E} \vee (F \times B)^{\infty}_\mathcal{E} \right) + 2q^2 \log k + qn_i \int f \, d\mu^{(i)}.
\]

Diving by \( n_i \), passing to the \( \lim \sup_{i \to \infty} \) and using the inequality 10 in [27], i.e.,

\[
H_{\mu} \left( \bigvee_{t=0}^{q-1} (\Theta^t)^{-1} Q|F_\mathcal{E} \vee (F \times B)^{\infty}_\mathcal{E} \right)
\]

19
\[
\geq \limsup_{n \to \infty} H_{\mu^{(\omega)}} \left( \bigvee_{t=0}^{q-1} (\Theta^t)^{-1} Q|_{\cal F_{\Theta}} \vee (\cal F \times \cal B)_{\Theta}^{-1} \right),
\]
we get
\[
qP_{\text{pre}}(T, f, \epsilon) \leq H_{\mu} \left( \bigvee_{t=0}^{q-1} (\Theta^t)^{-1} Q|_{\cal F_{\Theta}} \vee (\cal F \times \cal B)_{\Theta}^{-1} \right) + q \int f \, d\mu.
\]
Dividing by \( q \) and letting \( q \to \infty \), we have
\[
P_{\text{pre}}(T, f) \leq h_{\text{pre}, \mu}(T, Q) + \int f \, d\mu \leq h_{\text{pre}, \mu}(T) + \int f \, d\mu.
\]
Let \( \epsilon \to 0 \), then we have \( P_{\text{pre}}(T, f) \leq h_{\text{pre}, \mu}(T) + \int f \, d\mu \) and complete the proof of Theorem 8. \( \square \)

**Remark 9.** If \( f = 0 \), then, without any additional assumption, Theorem 8 can be expressed as \( h_{\text{pre}, \mu}(T) = \sup \{ h_{\text{pre}, \mu}(T) : \mu \in \cal M_{\text{P}}(\cal E, T) \} \). In [27], the variational principle for the pre-image topological entropy needs a measurability condition which in most cases cannot be satisfied. Thus Theorem 8 can be regarded as a revised version of Zhu’s. On the other hand, if \( \Omega \) consists of only one point, that is, \( \Omega = \{ \omega \} \), then by Remark 6, Theorem 8 generalizes Zeng’s deterministic variational principle on the whole space \( X \) [7] to any compact invariant subset \( E \).

**Acknowledgements**

The first author is supported by a grant from Postdoctoral Science Research Program of Jiangsu Province (0701049C). The second author is partially supported by the National Natural Science Foundation of China (10571086) and National Basic Research Program of China (973 Program) (2007CB814800).

**References**

[1] C. Castaing, M. Valadier, “Convex Analysis and Measurable Multifunctions, Vol. 580 of Lecture Notes in Mathematics,” Springer-Verlag, New York-Berlin, 1977.

[2] D. Ruelle, *Statistical mechanics on a compact set with zv action satisfying expansiveness and specification*, Trans. Amer. Math. Soc., 187 (1973), 237–251.
[3] D. Ruelle, “Thermodynamic Formalism,” Addison-Wesley, Reading, Mass, 1978.

[4] D. Fiebig, U. Fiebig, Z. Nitecki, *Entropy and preimage sets*, Ergodic Theory Dynam. Systems, **23** (2003), 1785–1806.

[5] Manfred Denker, Yuri Kifer, Manuel Stadlbauer, *Thermodynamic formalism for random countable Markov shifts*, Discrete Contin. Dyn. Syst., **22** (2008), 131–164.

[6] F. Ledrappier, P. Walters, *A relativised variational principle for continuous transformations*, J. London Math. Soc., **2** (1977), 568–576.

[7] F. Zeng, K. Yan, G. Zhang, *Pre-image pressure and invariant measures*, Ergodic Theory Dynam. Systems, **27** (2007), 1037–1052.

[8] K. Khanin, Y. Kifer, *Theorem formalism for random transformations and statistical mechanics*, in “Vol. 171 of Amer. Math. Soc. Transl. Ser. 2”, Amer. Math. Soc., Providence, 1996, 107–140.

[9] L. Arnold, “Random Dynamical Systems,” Springer-Verlag, New York, 1998.

[10] M. Misiurewicz, *A short proof of the variational principle for a $\mathbb{Z}_n$ action on a compact space*, Asterisque, **40** (1976), 147–157.

[11] M. Hurley, *On the topological entropy of maps*, Ergodic Theory Dynam. Systems, **15** (1995), 557–568.

[12] P. Walters, “An Introduction to Ergodic Theory,” Springer-Verlag, New York, 1982.

[13] R. Bowen, “Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms, Vol. 470 of Lecture Notes in Mathematics,” Springer-Verlag, 1975.

[14] R. Langevin, P. Walczak, *Entropie d’une dynamique*, C. R. Acad. Sci. Paris Sér. I Math., **312** (1991), 141–144.

[15] R. Langevin, F. Prezytycki, *Entropie de l’image inverse d’une application*, Bull. Soc. Math. France, **120** (1991), 237–250.
[16] R. M. Dudley, “Real Analysis and Probability,” Wadsworth and Brooks/Cole, Pacific Grove, CA, 1989.

[17] T. Bogenschütz, *Entropy, pressure, and a variational principle for random dynamical systems*, Random Comput. Dynam., 1 (1992), 99–116.

[18] S. Zelik, *On the Lyapunov dimension of cascade systems*, Comm. Pure Appl. Anal., 7 (2008), 971–985.

[19] T. Bogenschütz, “Equilibrium states for random dynamical systems,” Ph.D. thesis, Universität Bremen, 1993.

[20] T. Bogenschütz, V. M. Gundlach, *Ruelle’s transfer operator for random subshifts of finite type*, Ergodic Theory Dynam. Systems, 15 (1995), 413–447.

[21] W.-C. Cheng, S. Newhouse, *Pre-image entropy*, Ergodic Theory Dynam. Systems, 25 (2005), 1091–1113.

[22] W.-C. Cheng *Two-point pre-image entropy*, Discrete Contin. Dyn. Syst., 17 (2007), 107–119.

[23] Y. Pesin, “Dimension Theory in Dynamical Systems: Contemporary Views and Applications,” University of Chicago Press, Chicago and London, 1997.

[24] Y. Kifer, “Ergodic Theory of Random Transformations,” Birkhäuser, Boston, 1986.

[25] Y. Kifer, *Equilibrium states for random expanding transformations*, Random Comput. Dynam., 1 (1992), 1–31.

[26] Y. Kifer, *On the topological pressure for random bundle transformations*, in “Rokhlin’s Memorial Volume”, Amer. Math. Soc., 2001.

[27] Y. Zhu, *Preimage entropy for random dynamical systems*, Discrete Contin. Dyn. Syst., 18 (2007), 829–851.

[28] Z. Nitecki, F. Prezytycki, *Preimage entropy for mappings*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 9 (1999), 1815–1843.