A FORWARD ERGODIC CLOSING LEMMA AND THE
ENTROPY CONJECTURE FOR NONSINGULAR
ENDOMORPHISMS AWAY FROM TANGENCIES

SHUHEI HAYASHI

Graduate School of Mathematical Sciences
University of Tokyo, 3-8-1 Komaba, Tokyo, Japan

(Communicated by Jairo Bochi)

Abstract. We prove a forward Ergodic Closing Lemma for nonsingular $C^1$ endomorphisms, claiming that the set of eventually strongly closable points is a total probability set. The “forward” means that the closing perturbation is involved along a finite part of the forward orbit of a point in a total probability set, which is the same perturbation as in Mané’s Ergodic Closing Lemma for $C^1$ diffeomorphisms. As an application, Shub’s Entropy Conjecture for nonsingular $C^1$ endomorphisms away from homoclinic tangencies is proved, extending the result for $C^1$ diffeomorphisms by Liao, Viana and Yang.

1. Introduction. Let $M$ be a compact Riemannian manifold without boundary and let $\text{End}^1(M)$ be the space of $C^1$ endomorphisms of $M$ endowed with the $C^1$ topology. Denote by $\text{Diff}^1(M)$ the subset of $\text{End}^1(M)$ consisting of all $C^1$ diffeomorphisms of $M$. For $f \in \text{End}^1(M)$ denote by $\mathcal{M}_f(M)$ the set of $f$-invariant probability measures on the Borel $\sigma$-algebra of $M$, $\mathcal{M}_e(f)$ the ergodic elements of $\mathcal{M}_f(M)$ and $\text{Per}(f)$ the set of periodic points of $f$. For every neighborhood $U$ of $f \in \text{End}^1(M)$ and $\varepsilon > 0$ define a Borel set $\Sigma_+ (U, \varepsilon)$ as the set of points $x \in M$ such that there exist $g \in U$, $\ell \in \mathbb{Z}^+$ and $y \in \text{Per}(g)$ satisfying $g^\ell (y) = y$, $g = f$ on $M \setminus B^+_\varepsilon (f, x)$ and $d(f^j(x), g^j(y)) \leq \varepsilon$ for all $0 \leq j \leq \ell$, where $B^+_\varepsilon (f, x)$ is the closure of the $\varepsilon$-neighborhood of the forward orbit $O^+_f(x) = \{ f^j(x) : j \geq 0 \}$ of $x$, that is, the set of $w \in M$ such that $d(f^j(x), w) \leq \varepsilon$ for some $j \geq 0$. This means that $x \in \Sigma_+ (U, \varepsilon)$ is a forwardly $(U, \varepsilon)$-strongly closable point in the sense that the created periodic orbit for $g \in U$ by a forward perturbation $\varepsilon$-shadows a finite part of the forward $f$-orbit of $x$. Let $\Sigma_+ (U, f)$ be the set of $x$ belonging to $\Sigma_+ (U, \varepsilon)$ for every $\varepsilon > 0$, which can be written as

$$\Sigma_+ (U, f) = \bigcap_{n \geq 0} \Sigma_+ (U, \varepsilon_n)$$

for a sequence of positive numbers $\{ \varepsilon_n \}_{n \geq 0}$ converging to 0. Define

$$\Sigma^{ev}_+ (U, f) = \bigcup_{m \geq 0} f^{-m} (\Sigma_+ (U, f))$$

2010 Mathematics Subject Classification. 37C05, 37D30, 37B40, 37G25.

Key words and phrases. Ergodic closing lemma, entropy conjecture, upper semicontinuity of entropy, entropy expansiveness, homoclinic tangencies.
and then denote by $\Sigma_{+}^{\text{ev}}(f)$ the set of $x \in M$ belonging to $\Sigma_{+}^{\text{ev}}(U, f)$ for every neighborhood $U$ of $f$, which is the set of eventually strongly closable points. That is, if $x \in \Sigma_{+}^{\text{ev}}(f)$ then there exists $n_0 \geq 0$ such that $f^{m}(x)$ is forwardly $(U, \varepsilon)$-strongly closable for every neighborhood $U$ of $f$ and every $\varepsilon > 0$. If $\{U_n\}_{n \geq 0}$ is a basis of neighborhoods of $f$ in $\text{End}^{1}(M)$, we can write

$$
\Sigma_{+}^{\text{ev}}(f) = \bigcap_{n \geq 0} \Sigma_{+}^{\text{ev}}(U_n, f) = \bigcup_{m \geq 0} f^{-m} \left( \bigcap_{n \geq 0} \Sigma_{+}(U_n, f) \right).
$$

When $f \in \text{Diff}^{1}(M)$, we can define $\Sigma(U, \varepsilon)$ as the diffeomorphisms version of $\Sigma_{+}(U, \varepsilon)$ which was introduced by Mañé [14] who proved that given a neighborhood $U$ of $f \in \text{Diff}^{1}(M)$ and $\varepsilon > 0$, $\mu(\Sigma(U, \varepsilon)) = 1$ for every $\mu \in \mathcal{M}_{f}(M)$. This implies that $\bigcap_{n \geq 0} \Sigma_{+}(Y_n, \varepsilon_n)$ with $Y_n = U_n \cap \text{Diff}^{1}(M)$ is a total probability set of $f$, or a set having full $\mu$-measure for every $\mu \in \mathcal{M}_{f}(M)$. We consider an endomorphisms version of this. Unfortunately, we need to take the union of backward iterates of the set of forwardly strongly closable points to ensure the total probability even for nonsingular endomorphisms. However, it is still useful to obtain asymptotic properties of the forward orbits of points in a total probability set.

For $f \in \text{End}^{1}(M)$ and $p \in M$, we say that $p$ is a critical point of $f$ if $Df(T_pM)$ is not injective. Denote by $C(f)$ the set of critical points of $f$, which is a compact subset of $M$. When $C(f) = \emptyset$ we say that $f$ is nonsingular. Let $\text{NEnd}^{1}(M)$ be the set of all nonsingular endomorphisms in $\text{End}^{1}(M)$. Note that $\text{NEnd}^{1}(M)$ is an open subset of $\text{End}^{1}(M)$. The first theorem is a forward Ergodic Closing Lemma for endomorphisms without critical points on the closure of the union of all supports of $f$-invariant probability measures, denoted by

$$
S(f) = \{ x \in \text{supp}(\mu) : \mu \in \mathcal{M}_{f}(M) \}.
$$

**Theorem A.** Let $f \in \text{End}^{1}(M)$ satisfy $S(f) \cap C(f) = \emptyset$. Then

$$
\mu(\Sigma_{+}^{\text{ev}}(f)) = 1
$$

for every $\mu \in \mathcal{M}_{f}(M)$.

The “forward” means that the perturbation for the closing is involved along a finite part of the forward orbit of $\mu$-almost every $x$ for every $\mu \in \mathcal{M}_{f}(M)$. The “backward” Ergodic Closing Lemma for nonsingular $C^1$ endomorphisms has been proved by Moriyasu [19] (see also [5]). Its perturbation for the closing is based on Wen's $C^1$ Closing Lemma [30] for nonsingular endomorphisms, making the perturbation along a finite part of the backward orbit of a given nonwandering point. In the endomorphisms case, since prehistories of a given point is not unique, some different feature from the diffeomorphisms case may arise. One of the advantages of the forward Ergodic Closing Lemma is that we can safely understand some results for diffeomorphisms obtained by using Mañé’s Ergodic Closing Lemma still hold for nonsingular endomorphisms because the perturbation itself is the same.

For the statement of the next theorem, we need to give several definitions. Let $X$ be a compact metric space and let $f : X \to X$ be a continuous map. For $x \in X$, $\varepsilon > 0$ and $n \in \mathbb{Z}^{+}$, we denote by $B_n(x, \varepsilon)$ the dynamical ball (or Bowen ball) at $x$ of radius $\varepsilon > 0$ and length $n$, that is, the set of $y \in X$ such that $d(f^j(x), f^j(y)) \leq \varepsilon$ for all $0 \leq j < n$. Then, for a subset $K$ of $X$, an $(n, \varepsilon)$-spanning set $E$ for $K$ is a
set satisfying 
\[ K \subset \bigcup_{x \in E} B_n(x, \varepsilon). \]

Let \( r_n(K, \varepsilon) \) denote the smallest cardinality of any \((n, \varepsilon)\)-spanning set for \( K \). Then we also say that \( K \) is \((n, \varepsilon)\)-spanned by \( E \). Set 
\[ r(K, \varepsilon) = \limsup_{n \to +\infty} \frac{1}{n} \log r_n(K, \varepsilon). \]

In particular, we write 
\[ h(f, \varepsilon) = r(X, \varepsilon). \]

For a compact subset \( K \subset X \), let 
\[ h(f, K) = \lim_{\varepsilon \to 0} r(K, \varepsilon). \]

Then the topological entropy \( h(f) \) of \( f \) is defined by 
\[ h(f) = h(f, X). \]

Denote by 
\[ B_\infty(x, \varepsilon) \]
the set of \( y \in X \) such that 
\[ d(f^j(x), f^j(y)) \leq \varepsilon \]
for all \( j \geq 0 \). For \( \varepsilon > 0 \) and a subset \( A \) of \( X \), we say that \( f \) is \( \varepsilon \)-entropy expansive around \( A \) if 
\[ h(f, B_\infty(x, \varepsilon)) = 0 \]
for every \( x \in A \). When \( A = X \), we say that \( f \) is \( \varepsilon \)-entropy expansive. If there exists \( \varepsilon > 0 \) such that 
\[ \sup_{x \in A} h(f, B_\infty(x, \varepsilon)) = 0, \]
then \( f \) is said to be entropy expansive around \( A \). As an intermediate notion between one-sided expansiveness and entropy expansiveness, we say that \( f \) is linearly \( \epsilon \)-entropy expansive around \( A \) if 
\[ r_n(B_\infty(x, \varepsilon)) = 0 \]
for all \( n \) sufficiently large. It is easy to observe that this definition is independent of the choice of equivalent metric and therefore shared by topologically equivalent maps. When \( \varepsilon > 0 \) is specified for which \( f \) satisfies the inequality above, we say that \( f \) is linearly \( \varepsilon \)-entropy expansive around \( A \)

In dealing with \( f \in \text{End}^1(M) \), it is natural to consider the inverse limit space of \( f \) (see [23, 28] for instance). Define the inverse limit space \( M^f \) by 
\[ M^f = \{ \hat{x} = (x_i)_{i \leq 0} : x_i \in M, f(x_i) = x_{i+1} \text{ for all } i \leq -1 \}. \]

Moreover, for a compact subset \( \Lambda \subset M \) that is \( f \)-invariant (i.e., \( f(\Lambda) = \Lambda \)), define 
\[ \Lambda^f = \{ \hat{x} = (x_i)_{i \leq 0} : x_i \in \Lambda, f(x_i) = x_{i+1} \text{ for all } i \leq -1 \}. \]

Given \( \beta > 1 \), we endow a metric \( d_\beta \) on \( M^f \) by 
\[ d_\beta(\hat{x}, \hat{y}) = \sum_{i \leq 0} \beta^i d(x_i, y_i), \]
where \( d \) denotes the Riemannian metric on \( M \). It is easy to see that the topology on \( M^f \) with respect to this metric does not depend on the choice of \( \beta > 1 \), for which \( M^f \) and \( \Lambda^f \) are both compact sets. Let \( \pi : M^f \to M \) be the canonical projection defined by \( \pi(\hat{x}) = x_0 = x \) and denote by \( \hat{f} : M^f \to M^f \) the left shift
Then we have continuous families of $M$ groups of $\{\text{extending } [12, \text{Corollary C}] \}$ to the nonsingular endomorphisms case. The semicontinuity of the entropy map $\mu$ is well-known ([18]) that the entropy expansiveness implies the upper entropy expansiveness for nonsingular endomorphisms away from homoclinic tangencies. Let $\Lambda$ be a compact $f$-invariant set $\Lambda$ with $\Lambda \cap C(f) = \emptyset$ is hyperbolic if there exists a splitting $T_{\Lambda^f} = E^s \oplus E^u$ with $T_{\Lambda^f}(\hat{x}) = E^s(x_0) \oplus E^u(\hat{x})$ such that $Df(E^s(x)) = E^s(f(x))$ and $Dx_0 f(E^u(\hat{x})) = E^u(f(\hat{x}))$, and there are constants $C > 0$ and $0 < \lambda < 1$ satisfying $\|Df^n | E^s(x)\| \leq C\lambda^n$

and $m(Dx_0 f^n | E^u(\hat{x})) \geq C^{-1}\lambda^{-n}$ for all $x \in \Lambda$, $\hat{x} \in \Lambda^f$ and $n \geq 0$, where $m(T) = \min \{\|Tv\| : \|v\| = 1\}$ is the minimum norm for a linear transformation $T$. Then we say that $E^s$ and $E^u$ are contracting and expanding, respectively. In particular, it is called an Axiom A basic set if $\Lambda \subset \text{Per}(f)$, $f|\Lambda$ is transitive (i.e., $O^+_f(x)$ is dense in $\Lambda$ for some $x \in \Lambda$), and $\Lambda$ is isolated (i.e., $\bigcap_{j=0}^{+\infty} f^{-j}(U) = \Lambda$ for some open neighborhood $U$ of $\Lambda$). Then we have continuous families of $C^1$ embedded discs, local stable manifolds $\{W^s_k(x,f) : x \in \Lambda\}$ and local unstable manifolds $\{W^u_k(\hat{x},f) : \hat{x} \in \Lambda^f\}$ defined by:

$W^s_\delta(x,f) = \{y \in M : d(f^j(x), f^j(y)) \leq \delta \text{ for all } j \geq 0\}$,

$W^u_\delta(\hat{x},f) = \{y \in M : y \text{ has a prehistory } \hat{y} = (y_i)_{i \leq 0} \text{ such that } d(x_i, y_i) \leq \delta \text{ for all } i \leq 0\}$.

See [23, Theorem 2.1] for this fact and it is also known that if $y \in W^s_\delta(x,f)$ with $x \in \Lambda$ then $d(f^n(x), f^n(y)) \to 0$ as $n \to +\infty$, and if $y \in W^u_\delta(\hat{x},f)$ with $\hat{x} \in \Lambda^f$ then $y$ has a unique prehistory $\hat{y} = (y_i)_{i \leq 0}$ such that $y_i \in W^s_k(f^i(\hat{x}),f)$ for all $i \leq 0$ and $d(y_i, x_i) \to 0$ as $i \to -\infty$. For a hyperbolic periodic point $p \in \Lambda$ with period $m$, a homoclinic point associated with $p$ is a point in $W^s_\delta(p,f) \cap \bigcup_{j \geq 0} f^{-m+j}(W^u_\delta(\hat{p},f)) \setminus \{p\}$ for $\hat{p} = (p_i)_{i \leq 0}$ with $\pi(\hat{p}) = p$ such that $p_i \in O^+_f(p)$ for all $i \leq 0$. If there exists a nontransversal homoclinic point associated with some hyperbolic periodic point, we say that $f$ exhibits a homoclinic tangency. Denote by $HT$ the set of $f \in \text{End}^1(M)$ exhibiting a homoclinic tangency.

The second theorem concerns the robust entropy expansiveness for nonsingular $C^1$ endomorphisms away from homoclinic tangencies.

**Theorem B.** Let $f \in \text{NEnd}^1(M) \setminus HT$. Then there exist a neighborhood $U$ of $f$ and $\varepsilon > 0$ such that every $g \in U$ is $\varepsilon$-entropy expansive.

**Remark.** It is well-known ([18]) that the entropy expansiveness implies the upper semicontinuity of the entropy map $\mu \mapsto h_\mu(f)$ defined on the space of $f$-invariant probability measures. Consequently, as a corollary of Theorem B, we see that $f \in \text{NEnd}^1(M) \setminus HT$ has a measure of maximal entropy (see [29] for instance), extending [12, Corollary C] to the nonsingular endomorphisms case.

The last theorem is a contribution to Shub’s Entropy Conjecture. We denote by $f_{*,k} : H_k(M,\mathbb{R}) \to H_k(M,\mathbb{R})$, $0 \leq k \leq \dim M$, a linear action on the real homology groups of $M$ induced by a $C^1$ map $f : M \to M$. Set $\nu = \dim M$ and let $\text{sp}(f_{*}) = \max_{0 \leq k \leq \nu} \text{sp}(f_{*,k})$, where $\text{sp}(\cdot)$ denotes the spectral radius.
A FORWARD ERGODIC CLOSING LEMMA AND THE ENTROPY CONJECTURE 2289

where \( \text{sp}(f_{s,k}) \) is the spectral radius of \( f_{s,k} \). Shub’s Entropy Conjecture ([26]) claims that the logarithm of \( \text{sp}(f_s) \) is a lower bound for the topological entropy of \( f \):

\[
h(f) \geq \log \text{sp}(f_s).
\]

The full statement of the conjecture is known to be true when \( \dim M \leq 3 \) by the result of Manning [16] (together with Poincaré duality); i.e., \( \log \text{sp}(f_{s,1}) \leq h(f) \) holds for continuous maps. For the historical background, one can consult [8] or [12]. Here we focus on some recent progress on the Entropy Conjecture for \( f \in \text{Diff}^1(M) \). After the classical proofs (Shub and Williams [27], Ruelle and Sullivan [24]) of the Entropy Conjecture for Axiom A diffeomorphisms with no cycles in the early 70’s, relatively recently Saghin and Xia [25] proved it for partially hyperbolic diffeomorphisms with one-dimensional center bundle. Shortly afterwards, Liao, Viana and Yang [12] established the conjecture for \( f \in \text{Diff}^1(M) \) away from homoclinic tangencies. Since all the diffeomorphisms mentioned above are \( C^1 \) away from homoclinic tangencies, their result remarkably includes all of them. However, in the original view of the Entropy Conjecture, the invertibility of \( f \) is not essential, so its extension to \( C^1 \) endomorphisms should be considered.

By Yomdin’s influential work [32] that established the Entropy Conjecture for \( C^\infty \) maps, the proof of the conjecture for \( f \in \text{End}^1(M) \) is reduced to proving the upper semicontinuity of the topological entropy at \( f \). Actually, this approach has been used in [12] and other recent works ([3, 4, 13]). However, the \( C^1 \) creation of homoclinic tangencies, or the appearance of transversal homoclinic points by arbitrarily small \( C^1 \) perturbations can be the obstruction for its upper semicontinuity as observed in [17] by Misiurewicz. In fact, Liao, Viana and Yang [12] proved the upper semicontinuity at \( C^1 \) diffeomorphisms away from homoclinic tangencies through the robust entropy expansiveness. Our proof is in line with theirs. We first prove the one-sided version of [12, Proposition 2.5] for a continuous map \( f : X \to X \) on a compact metric space \( X \). From this it follows that the almost entropy expansiveness is sufficient to guarantee the full entropy expansiveness. Therefore, under a proposition claiming the robust linearly \( \varepsilon \)-entropy expansiveness around a total probability set we obtain the robust \( \varepsilon \)-entropy expansiveness as in Theorem B. For the proof of the proposition we use the forward Ergodic Closing Lemma by which similar arguments to the diffeomorphisms case become available.

For the robustness, we need \( S(g) \) having the \( \delta \)-neighborhood \( U_\delta(S(g)) \) of \( S(g) \) with some \( \delta > 0 \) independent of \( g \) such that \( U_\delta(S(g)) \cap C(g) = \emptyset \) for all \( g \) sufficiently close to \( f \in \text{End}^1(M) \setminus \overline{HT} \). However, it may happen that \( S(f_n) \) intersects an arbitrarily small neighborhood of \( C(f_n) \) as \( f_n \to f \) in the \( C^1 \) topology for some \( f_n \in \text{End}^1(M) \setminus \overline{HT} \), \( n \geq 1 \). This is the difficulty of extending Theorem B (and therefore Theorem C below) to \( f \in \text{End}^1(M) \setminus \overline{HT} \) satisfying only \( S(f) \cap C(f) = \emptyset \) that is the hypothesis of Theorem A.

As mentioned above, the Entropy Conjecture for \( f \in \text{NEnd}^1(M) \setminus \overline{HT} \) easily follows from the upper semicontinuity of the topological entropy at \( f \) through Yomdin’s theorem ([32]); i.e., the Entropy Conjecture is true for \( C^\infty \) maps on \( M \). In fact, choosing a convergent sequence of \( C^\infty \) maps \( f_n \to f \), \( n \geq 1 \), in the \( C^1 \) topology from the homotopy class of \( f \), we have \( \text{sp}((f_n)_*\alpha) = \text{sp}(f_*\alpha) \), and then

\[
h(f) \geq \limsup_{n \to +\infty} h(f_n) \geq \limsup_{n \to +\infty} \log \text{sp}((f_n)_*\alpha) = \log \text{sp}(f_*\alpha),
\]

where the upper semicontinuity is used in the first inequality. Similarly to the proof of diffeomorphisms case in [12], the robust \( \varepsilon \)-entropy expansiveness implies
the upper semicontinuity of the map $f \mapsto h(f)$ for $f \in \text{NEnd}^1(M) \setminus \text{HT}$ (Theorem 3.7), extending [12, Theorem D] and proving the Entropy Conjecture for nonsingular $C^1$ endomorphisms away from homoclinic tangencies.

**Theorem C.** If $f \in \text{NEnd}^1(M) \setminus \text{HT}$, then the Entropy Conjecture holds for $f$.

Let us give two typical examples of endomorphisms belonging to $\text{NEnd}^1(M) \setminus \text{HT}$. We say that a compact set $\Lambda \subset M \setminus C(f)$ with $f(\Lambda) = \Lambda$ admits a dominated splitting if there exists a continuous splitting $T_{\Lambda'} = E \oplus F$ with $T_{\Lambda'}(\hat{x}) = E(x_0) \oplus F(\hat{x})$ such that $D_xf(E(x)) = E(f(x))$ and $D_{x_0}f(F(\hat{x})) = F(\hat{f}(\hat{x}))$, and there is $m \in \mathbb{Z}^+$ satisfying

$$\|Df^m|E(x)\| \cdot m(D_{x_0}f^m|F(\hat{x}))^{-1} < 1/2$$

for all $x \in \Lambda$ and all $\hat{x} \in \Lambda'$ with $\pi(\hat{x}) = x_0 = x$. Then, we say that the splitting $T_{\Lambda'} = E \oplus F$ is $m$-dominated. More generally, when we have a splitting

$$T_{\Lambda'} = E^1 \oplus E^2 \oplus E^3$$

such that if $E^i \oplus E^j$ is dominated (resp. $m$-dominated) for any pair of nontrivial subbundles $E^i$ and $E^j$ with $1 \leq i < j \leq 3$, we say that the splitting is dominated (resp. $m$-dominated). If $M'$ admits a dominated splitting $T_{M'} = E^1 \oplus E^2 \oplus E^3$ with $\dim E^2 = 1$ such that $E^1$ and $E^3$ are contracting and expanding, respectively, $f$ is called partially hyperbolic with one-dimensional center, which is obviously $C^1$ far away from homoclinic tangencies. Thus, we see that nonsingular partially hyperbolic endomorphisms with one-dimensional center belong to $\text{NEnd}^1(M) \setminus \text{HT}$.

Theorems B and C will be proved in Section 3 after giving the Preliminaries in Section 2. Finally, we will prove Theorem A in Section 4.

2. **Preliminaries.** In this section, we give basic definitions and known facts that will be needed in the following sections. The first fact is the existence of locally invariant discs associated with dominated splittings. Let $\Lambda \subset M \setminus C(f)$ satisfy $f(\Lambda) = \Lambda$ and admit a dominated splitting

$$T_{\Lambda'} = E \oplus F \quad \text{with} \quad T_{\Lambda'}(\hat{x}) = E(x_0) \oplus F(\hat{x})$$
Applying the proof of [11, Theorem 5.5], we have a continuous family of $C^1$ discs \( \{ D_{\eta_0}^F(\hat{x}) : \hat{x} \in \Lambda' \} \) in \( M' \) tangent to \( F(\hat{x}) \) at \( \hat{x} \in M' \) with size \( \eta_0 > 0 \). These discs are locally invariant in the sense that there is a constant \( L_0 \geq 1 \) such that if \( 0 < \eta \leq \eta_0 \) then 
\[
\hat{f}(D_{\eta/L_0}(\hat{x})) \subset D_F^F(\hat{f}(\hat{x}))
\]
for all \( \hat{x} \in \Lambda' \). Moreover, by the same argument as in the proof of [10, Theorem 6.2], if \( \eta_0 > 0 \) is small enough, \( E(x) \), \( x \in \Lambda \), are extended to some \( \tilde{E}(y) \) over \( y \in \pi(D_{\eta_0}^F(\hat{x})) \), having a continuous family of \( C^1 \) discs \( \{ D_{\eta_0}^E(y) : y \in \pi(D_{\eta_0}^F(\hat{x})) \} \) in \( M \) tangent to \( \tilde{E}(y) \) at \( y \) with size \( \eta_0 > 0 \) and \( D_{\eta_0}^E(y) = D_{\eta_0}^E(x) \) when \( y = x \). These discs are also locally invariant in the sense that if \( 0 < \eta \leq \eta_0 \) then 
\[
\hat{f}(D_{\eta/L_0}(\hat{x})) \subset D_{\eta}(\hat{f}(\hat{x}))
\]
for all \( y \in \pi(D_{\eta_0}^F(\hat{x}) \cap \hat{f}^{-1}(D_{\eta_0}^F(\hat{f}(\hat{x})))) \) with any \( \hat{x} \in \Lambda' \).

Next we recall definitions and facts about topological entropy (see [29] for more details). The definition of the topological entropy given in the Introduction coincides with that of topological entropy using open covers. Let \( \alpha \) be a finite open cover of \( X \) and let \( |\alpha^n| \) be the smallest cardinality of a subcover of \( \bigvee_{j=0}^{n-1} f^{-j}\alpha \). Then 
\[
h(f) = \sup_{\alpha} h(f, \alpha) \quad \text{with} \quad \alpha \text{ ranging over all finite open covers, where}
\]
\[
h(f, \alpha) = \lim_{n \to +\infty} \frac{1}{n} \log |\alpha^n| = \inf_{n \geq 1} \frac{1}{n} \log |\alpha^n|.
\]
Define, for \( \varepsilon > 0 \), the \( \varepsilon \)-local entropy by
\[
h^\varepsilon(f, \alpha) = \sup_{x \in X} h(f, B_\infty(x, \varepsilon)).
\]
Bowen gave the following inequality in [2, Theorem 2.4]: given \( \varepsilon > 0 \),
\[
h(f) \leq h(f, \varepsilon) + h^\varepsilon(f, \varepsilon).
\](2.1)
In the proof of this inequality, Bowen used the following lemma, which will be also needed in the next section.

**Lemma 2.1 (Bowen [2, Lemma 2.1]).** Suppose \( 0 = t_0 < t_1 < \cdots < t_{r-1} < t_r = n \) and \( f^{+i}(K) \) is \((t_{i+1} - t_i, \varepsilon)\)-spanned by \( E_i \) for all \( 0 \leq i < r \). Then
\[
r_n(K, 2\varepsilon) \leq \prod_{0 \leq i < r} \#E_i.
\]
Since \( r_n(M, \varepsilon) \leq |\alpha^n| \) when \( \text{diam } \alpha \leq \varepsilon \) for every \( n \geq 1 \), we have \( h(f, \varepsilon) = r(X, \varepsilon) \leq h(f, \alpha) \), implying from (2.1) that
\[
h(f) \leq h(f, \alpha) + h^\varepsilon(f, \varepsilon)
\](2.2)
when \( \text{diam } \alpha \leq \varepsilon \). Given \( \varepsilon > 0 \), we say that a continuous map \( f : X \to X \) is \( \varepsilon \)-almost entropy expansive if there exists a total probability set \( A \subset X \) of \( f \) such that
\[
h(f, B_\infty(x, \varepsilon)) = 0
\]
for every \( x \in A \). This is the one-sided version of the definition introduced in [12] for homeomorphisms by Liao, Viana and Yang. A direct consequence of (2.2) shows that the upper semicontinuity of the topological entropy follows from the robust \( \varepsilon \)-entropy expansiveness with respect to the \( C^0 \) topology. In fact, when \( f_n \to f \) in the \( C^0 \) topology, letting \( \alpha \) be a finite open cover of \( X \) with \( \text{diam } \alpha \leq \varepsilon \), from (2.2)
together with the robust $\varepsilon$-entropy expansiveness and the upper semicontinuity of $\varepsilon \rightarrow h(f, \alpha)$, it follows that

$$\limsup_{n \rightarrow +\infty} h(f_n) \leq \limsup_{n \rightarrow +\infty} h(f_n, \alpha) + \limsup_{n \rightarrow +\infty} h^*(f_n, \varepsilon) \leq h(f, \alpha) \leq h(f).$$  \text{(2.3)}

Finally, we recall the so-called Pliss Lemma, which has been one of the basic tools in the study of weak hyperbolicity. Denote by $(x, y; f)$ with $y = f^n(x)$, $n \geq 0$, a finite set $\{f^j(x) : 0 \leq j \leq n\}$, which is called a string. For $f \in \text{End}^1(M)$ and $\Lambda \subset M \setminus C(f)$ with $f(\Lambda) = \Lambda$ admitting a dominated splitting $T_{\Lambda f} = E \oplus F$, we say that $(x, f^n(x); f)$ in $\Lambda$ is a $\gamma$-string over $E$ if

$$\prod_{j=0}^{n-1} \|Df| E(f^j(x))\| \leq \gamma^n.$$

In particular, if $(x, f^n(x); f)$ is a $\gamma$-string over $E$ for all $0 < j \leq n$, we call $(x, f^n(x); f)$ a uniform $\gamma$-string over $E$. Moreover, we say that $(\hat{x}, \hat{f}^n(\hat{x}); \hat{f})$ in $\Lambda^f$ with $\hat{x} = (x_i)_{i \leq 0}$ is a $\gamma$-string over $F$ if

$$\prod_{j=0}^{n-1} \|Df| E(\hat{f}^j(\hat{x}))\|^{-1} \leq \gamma^n.$$

In particular, if $(\hat{x}, \hat{f}^n(\hat{x}); \hat{f})$ is a $\gamma$-string over $F$ for all $0 \leq j < n$, we call $(\hat{x}, \hat{f}^n(\hat{x}); \hat{f})$ a uniform $\gamma$-string over $F$. The following lemma is due to Pliss [20] (see also [15, Lemma 11.8] and [12, Lemma 3.5]):

**Lemma 2.2 (Pliss Lemma).** For all $0 < \gamma < \bar{\gamma} < 1$ there exist $N(\gamma, \bar{\gamma}) > 0$ and $0 < c(\gamma, \bar{\gamma}) < 1$ such that the following properties hold.

(a) If $(x, f^n(x); f)$ is a $\gamma$-string over $E$ and $n \geq N(\gamma, \bar{\gamma})$, there exist $0 < n_1 < \cdots < n_k \leq n$, $k > 1$, such that $k \geq n c(\gamma, \bar{\gamma})$ and $(f^n(x), f^n(x); f)$ is a uniform $\bar{\gamma}$-string over $E$ for all $1 \leq i \leq k$.

(b) If $(\hat{x}, \hat{f}^n(\hat{x}); \hat{f})$ is a $\gamma$-string over $F$ and $n \geq N(\gamma, \bar{\gamma})$, there exist $0 < n_1 < \cdots < n_k \leq n$, $k > 1$, such that $k \geq n c(\gamma, \bar{\gamma})$ and $(\hat{x}, \hat{f}^n(\hat{x}); \hat{f})$ is a uniform $\bar{\gamma}$-string over $F$ for all $1 \leq i \leq k$.

3. **Proofs of Theorems B and C.** In [12], Liao, Viana and Yang proved that if $f \in \text{Diff}^1(M) \setminus \overline{HT}$, then there exist a neighborhood $U$ of $f$ and $\varepsilon > 0$ such that every $g \in U$ is $\varepsilon$-entropy expansive. More precisely, letting

$$B^\varepsilon_\infty(x, \varepsilon, g) = \{y \in M : d(g^j(x), g^j(y)) \leq \varepsilon \text{ for all } j \in \mathbb{Z}\},$$

they proved that if $f \in \text{Diff}^1(M) \setminus \overline{HT}$ then there exist a neighborhood $U$ of $f$, $\varepsilon > 0$ and a constant $C(\delta) > 0$ depending on an arbitrarily small $\delta > 0$ such that for all $g \in U$ and $n \geq 1$ sufficiently large,

$$r_n(B^\varepsilon_\infty(x, \varepsilon, g), \delta) \leq C(\delta)n$$

holds at every $x$ in a total probability set of $g$ (see the proof of [12, Theorem 3.1]). In this section, we extend this property to $g C^1$ close to $f \in \text{NEnd}^1(M) \setminus \overline{HT}$, replacing the property with $B^\varepsilon_\infty(x, \varepsilon, g)$ and $C(\delta)$ by $B_\infty(x, \varepsilon, g)$ and $C(x, \delta)$ at every $x$ in a total probability set of $g$, where

$$B_\infty(x, \varepsilon, g) = \{y \in M : d(g^j(x), g^j(y)) \leq \varepsilon \text{ for all } j \geq 0\}$$
and \( C(x, \delta) \) is a constant depending on \( x \) and \( \delta \). The following lemma, the one-sided version of [12, Proposition 2.5], achieves the \( \varepsilon \)-entropy expansiveness from the \( \varepsilon \)-almost expansiveness as \( a = 0 \). The proof is essentially the same as (or even slightly simpler than) that of [12, Proposition 2.5], but we give the proof for completeness.

**Lemma 3.1.** Let \( f : X \to X \) be a continuous map. Given \( a > 0 \), if there is a total probability set \( A \subset X \) of \( f \) such that \( h(f, B_\infty(x, \varepsilon)) \leq a \) for every \( x \in A \), then

\[
h(f, B_\infty(x, \varepsilon)) \leq a
\]

for every \( x \in X \).

**Proof.** Suppose on the contrary that \( h(f, B_\infty(x_0, \varepsilon)) > a \) for some \( x_0 \in X \). Let \( \delta > 0 \) be an arbitrarily small positive number. Then, we can take constants \( c > a_0 > a \) and a subsequence \( m_i \to +\infty \) of \( m = 1, 2, \ldots \) such that

\[
\frac{1}{m_i} \log r_{m_i}(B_\infty(x_0, \varepsilon), \delta) > c
\]

for all \( i \geq 1 \). Choosing a subsequence of \( m_i, i = 1, 2, \ldots \), if necessary, we may assume that a sequence of probabilities

\[
\mu_i = \frac{1}{m_i} \sum_{j=0}^{m_i-1} \delta^j f(x_0), \quad i \geq 1,
\]

converges in the weak* topology to some \( f \)-invariant probability measure \( \mu \) on the Borel \( \sigma \)-algebra of \( X \). Take \( b \in (a_0, c) \) and define \( \Gamma_n \subset X \) by the set of \( x \in X \) such that

\[
\frac{1}{m} \log r_m(B_\infty(x, \varepsilon), \delta/4) < b
\]

for all \( m \geq n \). Note that \( \Gamma_n \subset \Gamma_{n'} \) when \( n \leq n' \). From the hypothesis, it follows that

\[
\mu \left( \bigcup_{n \geq 1} \Gamma_n \right) = 1.
\]

By the regularity of \( \mu \), we can take a compact subset \( \Lambda_n \) of \( \Gamma_n \) with \( \mu(\Gamma_n \setminus \Lambda_n) \to 0 \) as \( n \to +\infty \). Fix some arbitrarily large \( n \geq 1 \) and choose an \((n, \delta/4)\)-spanning set \( E_n(y) \) for \( B_\infty(y, \varepsilon), y \in \Lambda_n \), with \( \#E_n(y) < e^{bn} \), which implies

\[
B_\infty(y, \varepsilon) \subset \bigcup_{z \in E_n(y)} B_n(z, \delta/4).
\]

Define a neighborhood

\[
U_n(y) = \text{Int} \bigcup_{z \in E_n(y)} B_n(z, \delta/2)
\]

of \( B_\infty(y, \varepsilon) \). Then, take an open neighborhood \( V_n(y) \) of \( y \in \Lambda_n \) such that if \( N_n(y) \in \mathbb{Z}^+ \) is large enough, we have

\[
B_{N_n(y)}(u, \varepsilon) \subset U_n(y)
\]

for every \( u \in V_n(y) \). Note that any subset of \( U_n(y) \) is \((n, \delta/2)\)-spanned by \( E_n(y) \).

Choose \( y_1, \ldots, y_s \) for which the compact set \( \Lambda_n \) is covered by \( V_n(y_j), 1 \leq j \leq s \), and set

\[
W_n = \bigcup_{j=1}^s V_n(y_j), \quad \text{and} \quad N_n = \max\{n, N_n(y_1), \ldots, N_n(y_s)\}.
\]
Here we may assume that \( \mu(\partial W_n) = 0 \) to have

\[
\lim_{i \to +\infty} \mu_i(W_n) = \mu(W_n),
\]

which can be arbitrarily close to 1. Let \( \kappa_\delta \in \mathbb{Z}^+ \) be the cardinality of a \( \delta/2 \)-dense subset of \( X \). Note that any subset of \( X \) is \((1, \delta/2)\)-spanned by the \( \delta/2 \)-dense subset. Since \( n \geq 1 \) can be arbitrarily large (depending on \( \delta \)), we can fix \( n \) so large that

\[
1 - \mu_i(W_n) < \frac{c - b}{2 \log \kappa_\delta}
\]

for all \( i \) sufficiently large. Now define \( 0 = t_0 < t_1 < \cdots < t_{k_i} = m_i \) (where we can suppose \( m_i > N_n \)) inductively as:

\[
t_{j+1} = \begin{cases} t_j + n, & \text{if } t_j \in T(n), \\ t_j + 1, & \text{otherwise}, \end{cases}
\]

where

\[
T(n) = \{ t \in [0, m_i - N_n] : f^t(x_0) \in W_n \}.
\]

When \( t_j \in T(n) \), choose \( s_j \in \{ 1, \ldots, s \} \) such that \( f^{s_j}(x_0) \in V_n(y_{s_j}) \), and then

\[
f^{s_j}(B_{m_i}(x_0, \varepsilon)) \subset B_{N_n}(f^{s_j}(x_0), \varepsilon) \subset U_n(y_{s_j}),
\]

which is \((n, \delta/2)\)-spanned by \( E_n(y_{s_j}) \). On the other hand, when \( t_j \notin T(n) \), we have \( f^{s_j}(B_{m_i}(x_0, \varepsilon)) \) that is \((1, \delta/2)\)-spanned by the \( \delta/2 \)-dense subset of \( X \). Set

\[
\nu(n) = m_i - n \# T(n) \geq 0.
\]

Then

\[
\nu(n) \leq \# \{ 0 \leq j < m_i : f^j(x_0) \notin W_n \} + N_n = (1 - \mu_i(W_n))m_i + N_n.
\]

Now apply Lemma 2.1 to obtain

\[
r_{m_i}(B_{m_i}(x_0, \varepsilon), \delta) \leq \kappa_\delta^{\nu(n)} \prod_{t_j \in T(n)} \# E_n(y_{s_j}) \leq \kappa_\delta^{\nu(n)} e^{b m_i \# T(n)}
\]

\[
\leq \kappa_\delta^{(1 - \mu_i(W_n))m_i + N_n} e^{b m_i}
\]

\[
= \exp\left( m_i \left( b + (1 - \mu_i(W_n)) \log \kappa_\delta + \frac{N_n}{m_i} \log \kappa_\delta \right) \right).
\]

Hence, using (3.2), we have

\[
\frac{1}{m_i} \log r_{m_i}(B_{m_i}(x_0, \varepsilon), \delta) \leq b + \frac{c - b}{2 \log \kappa_\delta} \log \kappa_\delta + \frac{N_n}{m_i} \log \kappa_\delta < c
\]

for large \( i \), contradicting (3.1). \( \square \)

Let \( f \in \text{End}^1(M) \) and let \( \Lambda \subset M \setminus C(f) \) be a compact set satisfying \( f(\Lambda) = \Lambda \). Then there is \( \varepsilon_0 > 0 \) such that \( f|_{B_{\varepsilon_0}(x)} \) is a diffeomorphism onto its image for all \( x \in \Lambda \), where \( B_r(x) = \{ y \in M : d(x, y) \leq r \} \). In what follows, we consider \( B_\infty(x, \varepsilon) \) as \( X = M \) for the Riemannian metric \( d \) on \( M \) with \( \dim M = \nu \). We can suppose \( \varepsilon > 0 \) of \( B_\infty(x, \varepsilon) \) is always smaller than \( \varepsilon_0 \). The weak hyperbolicity over the forward orbit of \( x \) defined below are important to prove Theorem B, which will be obtained through Theorem A.
For a dominated splitting $T_A = E \oplus F$, $\rho > 0$, $m \in \mathbb{Z}^+$ and $\hat{x} = (x_i)_{i \leq 0} \in \Lambda_f$, with $\pi(\hat{x}) = x_0 = x$, we say that $E$ is $(\rho, m, f)$-contracting at $\hat{x}$ if

$$\limsup_{\ell \to +\infty} \frac{1}{\ell} \sum_{j=0}^{\ell-1} \log \left\| Df^m | E(f^{m_j}(x)) \right\| < -\rho$$

and $F$ is weakly $(\rho, m, f)$-expanding at $\hat{x}$ if

$$\liminf_{\ell \to +\infty} \frac{1}{\ell} \sum_{j=0}^{\ell-1} \log \left( Df^{m_j}(x) \right) F(f^{m_j}(\hat{x})) < -\rho.$$

Let $T_A = E^1 \oplus E^2 \oplus E^3$ with $\dim E^2 \leq 1$ be a dominated splitting. We put subbundles $E^s, E^{cs}, E^{cu}$ and $E^u$ as:

$$E^s = E^1, \quad E^{cs} = E^1 \oplus E^2, \quad E^{cu} = E^2 \oplus E^3 \quad \text{and} \quad E^u = E^3.$$

Now we introduce a coordinatelike structure in $B_{\infty}(x, \varepsilon)$ and its positive iterates for $x \in \Lambda$ and small $\varepsilon > 0$ by using locally invariant $C^1$ discs given in Section 2. For small $\eta > 0$ let $\hat{D}_\eta^\sigma(\hat{x})$ and $D_\eta^\sigma(y)$ with $y \in \pi(\hat{D}_\eta^\sigma(\hat{x}))$, $(\sigma, \sigma') \in \{(s, u), (s, cu), (cs, u)\}$, be locally invariant $C^1$ discs tangent to $E^\sigma(\hat{x})$ and $E^{\sigma'}(y)$ at $\hat{x}$ and $y$, respectively, with size $\eta > 0$ given in Section 2 for a dominated splitting

$$T_A = E^\sigma \oplus E^{\sigma'}.$$

Here we suppose $0 < \eta \leq \eta_0$ and $0 < \varepsilon \leq \eta/(3L_0)$ for $\eta_0 < \varepsilon_0/2$ and $L_0$ given in Section 2 as $E \oplus F = E^\sigma \oplus E^{\sigma'}$. For every $x \in \Lambda$ and every $\hat{x} \in \Lambda_f$ with $\pi(\hat{x}) = x$, define

$$D_\eta^\sigma(\hat{x}) = \pi(\hat{D}_\eta^\sigma(\hat{x})).$$

Then

$$f(D_\eta^\sigma(L_0/\hat{x})) \cap \pi(\hat{D}_\eta^\sigma(\hat{x})) = D_\eta^\sigma(f(\hat{x})).$$

For every $\hat{x} \in \Lambda_f$, define

$$\mathcal{F}_\eta(\hat{x}) = D_\eta^\sigma(\hat{x}), \quad \sigma' \in \{cu, u\},$$

and let

$$\mathcal{F}_\eta = \{\mathcal{F}_\eta(\hat{x}) : \hat{x} \in \Lambda_f\}$$

be a locally invariant $C^1$ disc family over $\Lambda_f$. Next consider $D_\eta^\sigma(x^u)$ with $x^u \in \mathcal{F}_\eta(\hat{x}), \sigma \in \{s, cs\}$, satisfying

$$f(D_\eta^\sigma(L_0/\hat{x})) \cap \mathcal{F}_\eta(f(x^u)),$$

which is independent of the choice of $\hat{x}^u$ with $\pi(\hat{x}^u) = x^u$.

Now, for $\sigma \in \{s, cs\}$ we have a continuous locally invariant $C^1$ disc family

$$\mathcal{F}_\sigma = \{\mathcal{F}_\sigma(\hat{x}) : x^u \in \mathcal{F}_\eta(\hat{x})\}, \quad \hat{x} \in \Lambda_f,$$

with a pre-lamination structure (see [11]) whose leaves $\mathcal{F}_\sigma(x^u)$ at $x^u \in \mathcal{F}_\eta(\hat{x})$ are defined by

$$\mathcal{F}_\sigma(x^u) = D_\eta^\sigma(x^u), \quad \sigma \in \{s, cs\}.$$
\( \pi(\hat{x}) = x \), by the continuity of \( \mathcal{F}_\varepsilon^u \), it is easy to see that \( \bigcup_{x^u \in \mathcal{F}^u(\hat{x})} \mathcal{F}_\varepsilon^u(x^u) \) contains a neighborhood of \( x \) in \( M \) by the same argument as in the proof of [10, Lemma 4.1]. When \( E^s \) is trivial and \( T_{\Lambda^f} = E^s \oplus E^u \), similarly \( \bigcup_{x^u \in \mathcal{F}^u(\hat{x})} \mathcal{F}_\varepsilon^s(x^u) \) contains a neighborhood of \( x \) in \( M \). We consider both cases simultaneously by assuming \( \sigma = cs \) in the former case and \( \sigma = s \) in the latter case. By the uniformity with respect to \( x \in \Lambda \) and \( \hat{x} \in \Lambda^f \) of \( x^u \mapsto \mathcal{F}_\varepsilon^u(\hat{x}) \) as a continuous section over \( \mathcal{F}^u(\hat{x}) \) to the space of \( C^1 \) embeddings of the closed disc with dimension \( \dim \mathcal{F}^\sigma(x) \) into \( M \), we see that

\[
B_\varepsilon(x) \subset \bigcup_{x^u \in \mathcal{F}^u(\hat{x})} \mathcal{F}_\varepsilon^u(x^u)
\]

for some \( \varepsilon > 0 \) independent of the choice of \( x \in \Lambda \) and \( \hat{x} \in \Lambda^f \). Then suppose \((\varepsilon_0, > \varepsilon_0 > 0 \) is so small that \( \varepsilon < \varepsilon_0 \) and if \( y \in \mathcal{F}_\varepsilon^u(\hat{x}) \cap B_\varepsilon(x) \) then

\[
\mathcal{F}^u(\hat{x}) \cap \mathcal{F}_\varepsilon^u(x^u) \subset B_{\varepsilon_0}(x).
\]

For every \( y \in B_{\infty}(x, \varepsilon) \) with \( \varepsilon > 0 \) as above there exists \( x^u \in \mathcal{F}^u(\hat{x}) \) such that \( y \in \mathcal{F}_\varepsilon^u(x^u) \), and then we call \( d(y, x^u) \) an \( s \)-distance of \( y \) from \( x \) and \( d(x^u, x) \) a \( u \)-distance of \( y \) from \( x \). By local invariance of disc families \( \mathcal{F}^u \) and \( \mathcal{F}_\varepsilon^u \), if \( y \in \mathcal{F}_\varepsilon^u(x^u) \cap B_{\infty}(x, \varepsilon) \) with \( x^u \in \mathcal{F}^u(\hat{x}) \) then for any fixed \( j \geq 0 \) we have \( f^j(x^u) \) converging to \( f^j(x) \) as \( \varepsilon \to 0 \) uniformly over \( x \in \Lambda, \hat{x} \in \Lambda^f \) and \( j \geq 0 \). This implies that \( T_{\omega} \mathcal{F}_\varepsilon^u(f^j(x^u)) \) with \( w \in \text{Int} \mathcal{F}_\varepsilon^u(f^j(x^u)) \cap B_{\text{d}(f^j(y), f^j(x^u))}(f^j(x^u)) \) converges to \( E^\sigma(f^j(x)) \) as \( \varepsilon \to 0 \) uniformly over \( x \in \Lambda, \hat{x} \in \Lambda^f \) and \( j \geq 0 \), where \( \hat{x}_j = f^j(\hat{x}) \). Therefore, for every \( \theta > 0 \) there exists \( 0 < \varepsilon_\theta < \varepsilon \) independent of \( x \in \Lambda, \hat{x} \in \Lambda^f \) and \( j \geq 0 \) such that if \( y \in \mathcal{F}_\varepsilon^u(x^u) \cap B_{\infty}(x, \varepsilon_\theta) \) with \( x^u \in \mathcal{F}^u(\hat{x}) \) then

\[
T_{\omega} \mathcal{F}_\varepsilon^u(f^j(x^u)) \subset C_{\theta}^0(\hat{x}_j) + w
\]

for any \( w \in \text{Int} \mathcal{F}_\varepsilon^u(f^j(x^u)) \cap B_{\text{d}(f^j(y), f^j(x^u))}(f^j(x^u)) \). By this uniform structure, making \( \varepsilon > 0 \) smaller if necessary we can suppose that

\[
B_\varepsilon(f^j(x)) \subset \bigcup_{x^u \in \mathcal{F}_\varepsilon^u(\hat{x})} \mathcal{F}_\varepsilon^u(x^u)
\]

for all \( x \in \Lambda, \hat{x} \in \Lambda^f \) and \( j \geq 0 \). Take \( 0 < \varepsilon < \varepsilon_\theta < \varepsilon \) with small \( 0 < \theta < 1 \). Then the properties of the derivatives on \( E^\sigma \) over strings \((x, f^m(x); f), m \geq 1 \), are carried over to those on tangent spaces of a part of locally invariant discs \( \mathcal{F}_\varepsilon^u(x^u) \), \( 0 \leq j \leq n \), with any fixed \( \hat{x} \in \Lambda^f \). As \( n \to +\infty \), we can understand the dynamics of the \( \varepsilon \)-neighborhood of the forward orbit of \( x \) through the coordinate-like structure around \( f^j(x) \), \( j \geq 0 \). Indeed, for every \( y \in B_{\infty}(x, \varepsilon) \) if an \( s \)-distance and a \( u \)-distance of \( y \) from \( x \) are defined, then for any \( j \geq 0 \) those of \( f^j(y) \) from \( f^j(x) \) are also defined as the coordinate-like structure around \( f^j(x) \). Moreover, not only \( j = 0 \) we can suppose if \( y \in B_{\infty}(x, \varepsilon) \) has \( x^u \) as in \( (3.3) \) then

\[
\mathcal{F}_\varepsilon^u(x^u) \cup \mathcal{F}_\varepsilon^u(f^j(x^u)) \subset B_{\varepsilon_0}(f^j(x))
\]

for all \( j \geq 0 \), restricted to which \( f \) is a diffeomorphism onto its image keeping the coordinate-like structure as above.

For the proofs of the following three lemmas, we take \( e^{-\rho} < \gamma < 1 \) for \( \rho > 0 \) given in the lemmas and let \( 0 < \varepsilon < \eta/(3L_0) \) be so small that if \( \varepsilon < \varepsilon \) then \( \exp_{f^j(x)}^{-1}(\mathcal{F}_\varepsilon^u(x^u)) \cap B_{\varepsilon_0}(f^j(x)) \cap \mathcal{F}_\varepsilon^u(f^j(x)) \) with any \( x^u \in \mathcal{F}_\varepsilon^u(\hat{x}) \) well approximate \( E^u(\hat{x}_j) \) and \( D_{f^j(x^u)} \exp_{f^j(x)}^{-1}(\tilde{E}^\sigma(f^j(x^u))) \) in the \( \varepsilon \)-ball of \( T_{f^j(x)}M \) for all \( j \geq 0 \).
Lemma 3.2. For a dominated splitting $T_M = E \oplus F$ with $\dim F = 1$. Suppose that there exist $m \in \mathbb{Z}^+$ and $\rho > 0$ such that $E$ is $(\rho, m, f)$-contracting at $\hat{x} \in M$. Then, there exist $\varepsilon > 0$ and $\ell_1 \in \mathbb{Z}^+$ such that given $\delta > 0$ arbitrarily small,

$$r_n(B_{\infty}(x, \varepsilon), 2\delta) \leq \kappa^{m\ell_1}(3\varepsilon/\delta)^{\nu}n$$

for all $n > m\ell_1$, where $\kappa \geq 1$ is a constant depending only on $\delta$.

**Proof.** In this case, we write $E^s = E$ and $E^u = F$. Since $E^s$ is $(\rho, m, f)$-contracting at $\hat{x}$, we have

$$\lim \sup_{n \to +\infty} \frac{1}{[n/m]} \sum_{j=0}^{[n/m]-1} \log \|Df^m|E^s(f^mj(x))\| < -\rho.$$ 

From this and Lemma 2.2 (a) (with $f$, $\gamma$ and $\tilde{\gamma}$ in Lemma 2.2 replaced by $f^m$, $e^{-\rho}$ and $\gamma$, respectively) there exists $0 < \ell_1 < [n/m]$ such that $(f^m, f^m[n/m])$ is a uniform $\gamma$-string over $E^s$ for large $n \geq 1$. Suppose that $\ell_1 > 0$ is the smallest positive integer given by Lemma 2.2 (a) satisfying this property. Then, by (3.3) and (3.4) with $s = \sigma$, for $0 < \theta(\gamma) < 1$ chosen small enough there exists $0 < \varepsilon < \varepsilon_\theta(\gamma) < \bar{\varepsilon}$ such that

$$d(f^mj(x^u), f^mj(y)) \leq \bar{\varepsilon} \sqrt{\gamma}^{1-\ell_1}$$

for all $\ell_1 \leq j \leq [n/m]$ and $y \in \mathcal{F}_x^s(x^u) \cap B_{\infty}(x, \varepsilon)$ with $x^u \in \mathcal{F}^u(\hat{x})$. That is, the $s$-distance of $f^mj(y)$ from $f^mj(x)$ decreases exponentially as $\ell_1 \leq j \to +\infty$. Since $\mathcal{F}_x^s(f^mj(x^u))$ is $(\nu - 1)$-dimensional, if $\varepsilon > 0$ is small enough and $n > m\ell_1$ then the smallest cardinality of $(n - m\ell_1, \delta)$-spanning set for $\mathcal{F}_x^s(f^mj(x^u)) \cap B_{\infty}(f^mj(x), \varepsilon)$ is $\approx (3\varepsilon/\delta)^{\nu-1}$ for all $x^u \in \mathcal{F}^u(\hat{x})$. Therefore, letting $\kappa \in \mathbb{Z}^+$ be the cardinality of a $\delta$-dense subset of $M$ and using Lemma 2.1, we obtain

$$r_n(\mathcal{F}_x^s(x^u) \cap B_{\infty}(x, \varepsilon), 2\delta) \leq \kappa^{m\ell_1}(3\varepsilon/\delta)^{\nu-1} \quad (3.5)$$

for all $x^u \in \mathcal{F}^u(\hat{x})$ and $n > m\ell_1$. On the other hand, since $\mathcal{F}^u(\hat{x})$ is one-dimensional, we have

$$r_n(\mathcal{F}^u(\hat{x}) \cap B_{\infty}(x, \varepsilon), \delta) \leq (3\varepsilon/\delta)n$$

for all $n \geq 1$. Thus, making $\varepsilon = \varepsilon(\theta(\gamma)) > 0$ smaller if necessary, the coordinateline structure and these properties give

$$r_n(B_{\infty}(x, \varepsilon), 2\delta) \leq \kappa^{m\ell_1}(3\varepsilon/\delta)^{\nu-1}. (3\varepsilon/\delta)n \leq \kappa^{m\ell_1}(3\varepsilon/\delta)^{\nu}n$$

for all $n > m\ell_1$. \(\square\)

Lemma 3.3. For a dominated splitting $T_M = E \oplus F$ with $\dim F = 1$. Suppose that there exist $m \in \mathbb{Z}^+$ and $\rho > 0$ such that $F$ is weakly $(\rho, m, f)$-expanding at $\hat{x} \in M$. Then, there exists $\varepsilon > 0$ such that given $\delta > 0$ arbitrarily small,

$$r_n(B_{\infty}(x, \varepsilon), \delta) \leq (3\varepsilon/\delta)n$$

for all $n \geq 1$.

**Proof.** In this case, we write $E^s = E$ and $E^u = F$. Since $F$ is weakly $(\rho, m, f)$-expanding at $\hat{x}$, we have

$$\lim \inf_{n \to +\infty} \frac{1}{[n/m]} \sum_{j=0}^{[n/m]-1} \log m(Df^mj(x)f^m|E^u(f^mj(\hat{x})))^{-1} < -\rho.$$
2298

SHUHEI HAYASHI

From this and Lemma 2.2 (b), there exist a subsequence \( n_k \to +\infty \) and \( 1 \leq \ell_k \leq [n_k/m] \) with \( \ell_k \to +\infty \) such that \( (\hat{x}, f^{m\ell_k}(\hat{x}); f^m) \) is a uniform \( \gamma \)-string over \( E^\nu \).

For \( 0 \leq j < \ell_k \) set

\[
g_j = f^{m(\ell_k-j)}|_{B_m(\ell_k-j)+1}(f^{m\ell}(x), \varepsilon_0),
\]

which is a diffeomorphism onto its image. If \( 0 < \varepsilon < \varepsilon_{\theta(\gamma)} < \varepsilon \) has been chosen small enough, by (3.4) with \( \sigma = s \), we have

\[
d(g_j^{-1}(z_k), g_j^{-1}(\hat{z}_k)) \leq \varepsilon / \gamma^{\ell_k-j}
\]

for all \( z_k = \pi(x_m\ell_k) \) and \( \hat{z}_k \in F^u(x_m\ell_k) \cap f^{m\ell_k}(B_\infty(x, \varepsilon)) \). Then the maximal \( u \)-distance of \( y_k = \theta_0^{-1}(\hat{z}_k) \) from \( x \) (ranging over all such \( \hat{z}_k \)) is arbitrarily small for large \( k \). Since

\[
B_\varepsilon(f^j(x)) \subset \bigcup_{x'_j \in F^u(\hat{x}_j)} \mathcal{F}_{\hat{x}_j}^u(x'_j)
\]

for all \( 0 \leq j \leq m\ell_k \), this implies that any \( y \notin \mathcal{F}_{\hat{x}}^u(x) \cap B_\varepsilon(x) \) cannot belong to \( B_\infty(x, \varepsilon) \) because then \( f^j(y) \notin B_\varepsilon(f^i(x)) \) for some \( 0 \leq j \leq m\ell_k \) with large \( k \). Therefore

\[
r_n(B_\infty(x, \varepsilon), \delta) = r_n(\mathcal{F}_{\hat{x}}^u(x) \cap B_\infty(x, \varepsilon), \delta)
\]

for all \( n \geq 1 \). On the other hand, since \( \mathcal{F}_{\hat{x}}^u(x) \) is one-dimensional, we have

\[
r_n(\mathcal{F}_{\hat{x}}^u(x) \cap B_\infty(x, \varepsilon), \delta) \leq (3\varepsilon/\delta)n \quad \text{(3.7)}
\]

for all \( n \geq 1 \). Combining (3.6) and (3.7), we obtain the required inequality. \( \square \)

Lemma 3.4. For a dominated splitting \( T_{\Lambda_f} = E^1 \oplus E^2 \oplus E^3 \) with \( \dim E^2 \leq 1 \).

Suppose that there exist \( m \in \mathbb{Z}^+ \) and \( \rho > 0 \) satisfying the following properties:

(a) \( E^1 \) is \( (\rho, m, f) \)-contracting at \( \hat{x} \in \Lambda_f \);

(b) \( E^3 \) is weakly \( (\rho, m, f) \)-expanding at \( \hat{x} \in \Lambda_f \).

Then, there exist \( \varepsilon > 0 \) and \( \ell_1 \in \mathbb{Z}^+ \) such that given \( \delta > 0 \) arbitrarily small,

\[
r_n(B_\infty(x, \varepsilon), 2\delta) \leq \kappa^{m\ell_1}(3\varepsilon/\delta)^\nu n
\]

for all \( n > m\ell_1 \), where \( \kappa \geq 1 \) is a constant depending only on \( \ell_1 \).

Proof. The proof is given combining arguments for the proofs of the previous two lemmas. Let \( \nu_1 = \dim E^s < \nu = \dim M \). Then, if \( 0 < \varepsilon < \varepsilon_{\theta(\gamma)} < \varepsilon \) is small enough then property (a) of the hypothesis and Lemma 2.2 (a) imply that there exists \( \ell_1 \geq 1 \) such that

\[
r_n(\mathcal{F}_{\hat{x}}^u(x^{cu}) \cap B_\infty(x, \varepsilon), 2\delta) \leq \kappa^{m\ell_1}(3\varepsilon/\delta)^\nu_1
\]

for all \( x^{cu} \in F^{cu}(\hat{x}) \) and \( n > m\ell_1 \), which corresponds to (3.5) and the proof is similar.

On the other hand, from property (b) of the hypothesis and Lemma 2.2 (b) there exists a sequence \( \ell_k \to +\infty \) such that if \( y_k \in B_\infty(x, \varepsilon) \) is written by \( f^{m\ell_k}(y_k) = w_k \) for some \( w_k \in \mathcal{F}_{\hat{x}m\ell_k}^u(f^{m\ell_k}(x^{cu})) \) with \( x^{cu} \in F^u(\hat{x}) \), then the maximal \( u \)-distance of such \( y_k \) from \( x \) is arbitrarily small for large \( k \), which is independent of the dimension of \( E^2 \) that is 0 or 1. Then we have

\[
r_n(B_\infty(x, \varepsilon), \delta) = r_n(\mathcal{F}_{\hat{x}}^u(x) \cap B_\infty(x, \varepsilon), \delta)
\]

for all \( n \geq 1 \), which corresponds to (3.6) and the proof is similar. Since the dimension of \( \mathcal{F}_{\hat{x}}^u(x) \cap F^{cu}(\hat{x}) \) is \( \leq 1 \), we have

\[
r_n(\mathcal{F}_{\hat{x}}^u(x) \cap F^{cu}(\hat{x}) \cap B_\infty(x, \varepsilon), \delta) \leq (3\varepsilon/\delta)n
\]

(3.10)
for all $n \geq 1$. Thus, making $\varepsilon = \varepsilon(\theta(\gamma)) > 0$ smaller if necessary, the coordinatelike structure and properties (3.8), (3.9) and (3.10) give
\[
r_n(B_\infty(x, \varepsilon), 2\delta) \leq \kappa^{m\ell_1}(3\varepsilon/\delta)^{\kappa n} \cdot (3\varepsilon/\delta)^n \leq \kappa^{m\ell_1}(3\varepsilon/\delta)^n n
\]
for all $n > m\ell_1$.

When $n \to +\infty$ in these lemmas, the only concern is that whether $\ell_1 = \ell_1(x, n)$, $n \geq 1$, in the proofs of Lemmas 3.2 and 3.4 are bounded independent of the choice of $n$ or not. If there is a sequence $n_j \to +\infty$ for which $\ell_1(x, n_j)$ is not bounded, setting $t_j = \ell_1(x, n_j)$, we have a sequence of strings $(x, f^{m_{t_j}}(x); f^m)$, $j \geq 1$, such that $(x, f^{m_{t_j}}(x); f^m)$ with large $j$ is not an $e^{-p}$-string over $E^*$. In fact, otherwise applying Lemma 2.2 (a) to the string $(x, f^{m_{t_j}}(x); f)$, we have $t_j > k_j \geq 0$ such that $(f^{\kappa_{t_j}}(x), f^{m_{t_j}}(x); f^m)$ is a uniform $\gamma$-string over $E^*$. Then $(f^{\kappa_{t_j}}(x), f^{m_{t_j}}(x); f^m)$ is also a uniform $\gamma$-string over $E^*$ with $k_j$ appeared before $t_j = \ell_1(x, n_j)$, satisfying the property of $\ell_1(x, n_j)$ for large $j$ and contradicting the smallest choice of $\ell_1(x, n_j)$. Thus $E^*$ in the proofs of Lemma 3.2 and 3.4 is not $(\rho, m, f)$-contracting at $\hat{x}$, which contradicts our hypothesis and we conclude
\[
\sup_{n} \ell_1(x, n) < +\infty.
\]

(3.11)

Now let us consider the following cases for a dominated splitting $T_{\Lambda^f} = E^1_f \oplus E^2_f \oplus E^3_f$ with dim $E^3_f \leq 1$ and a total probability set $\Lambda_0 \subset \Lambda$ of $f$: for every $\hat{x} \in \Lambda^f$ with $\pi(\hat{x}) = x \in \Lambda_0$,

(i) $E^1_f$ is trivial and $E^2_f$ is $(\rho, m, f)$-contracting at $\hat{x}$.

(ii) $E^1_f$ is trivial and $E^2_f$ is weakly $(\rho, m, f)$-expanding at $\hat{x}$.

(iii) $E^2_f$ is $(\rho, m, f)$-contracting and $E^3_f$ is weakly $(\rho, m, f)$-expanding at $\hat{x}$.

By (3.11) and Lemmas 3.2, 3.3 and 3.4, in all cases (i)-(iii) we have $0 < \theta(\gamma) < 1$ and $\varepsilon = \varepsilon(\theta(\gamma)) > 0$ with $e^{p} < \gamma < 1$ for which there exists a constant $C(x, \delta) > 0$ such that given $\delta > 0$ arbitrarily small,
\[
r_n(B_\infty(x, \varepsilon), 2\delta) \leq C(x, \delta)n
\]
for all $x \in \Lambda_0$ and $n$ sufficiently large.

From this together with the following lemma, we obtain the robust $\varepsilon$-almost entropy expansiveness.

**Lemma 3.5.** Let $f \in \text{NEnd}^1(M) \setminus \overline{HT}$. There exist a neighborhood $U_0$ of $f$, $m_0 \in \mathbb{Z}^+$ and $\rho_0 > 0$ such that for every $g \in U_0$ we have an $m_0$-dominated splitting
\[
T_{S(g)^g} = E^1_g \oplus E^2_g \oplus E^3_g
\]
with dim $E^2_g \leq 1$ and a total probability set $\Lambda_0(g)$ of $g$ such that $E^1_g$ is $(\rho_0, m_0, g)$-contracting and $E^3_g$ is weakly $(\rho_0, m_0, g)$-expanding at $\hat{x} \in S(g)^g$ with $\pi(\hat{x}) = x \in \Lambda_0(g)$.

Apply (3.12) to $f = g \in U_0$, $\Lambda = S(g)$, $\Lambda_0 = \Lambda_0(g)$, $m = m_0$ and $\rho = \rho_0$ given by Lemma 3.5. By the uniform choice of $U_0$, $m_0$ and $\rho_0$ in Lemma 3.5, it is easy to see that $\varepsilon > 0$ in (3.12) including the choice of $\varepsilon < \varepsilon < \varepsilon_0$ (guaranteeing a coordinatelike structure in the $\varepsilon$-neighborhood of $x$ around which $g$ is a diffeomorphism onto its image) can be fixed independent of the choice of $x \in \Lambda_0(g)$ and $g \in U_0$, where we may suppose $U_0 \subset \text{NEnd}^1(M) \setminus \overline{HT}$. Thus we have proved the following proposition.
Proposition 3.6. Let $f \in \text{NEnd}^1(M) \setminus \overline{HT}$. Then there exist a neighborhood $U_0$ of $f$ and $\varepsilon > 0$ such that every $g \in U_0$ is linearly $\varepsilon$-entropy expansive around a total probability set of $g$.

The diffeomorphisms version of Lemma 3.5 is included in Crovisier [6, Corollary 1.3] and Liao, Viana and Yang [12, Proposition 3.4]. These results should be extended to our case just replacing $\|Dg^{-m}\|E_g^1(g^m(\tau/k)(x)))$, $j \geq 0$, by

$$m(Dg^{-m}\|E_g^1(g^m(\tau/k)(x)))^{-1}$$

for some $m \in \mathbb{Z}^+$ using Theorem A. In fact, by hypothesis, if $f \in \text{NEnd}^1(M) \setminus \overline{HT}$ then there exist a neighborhood $U_0$ of $f$ and $\varepsilon_0 > 0$ such that for every $g \in U_0$ the restriction of $g \in U_0$ to the $\varepsilon_0$-neighborhood of any point of $M$ is a diffeomorphism onto its image. Therefore, as long as perturbations are made in a sufficiently small neighborhood of a periodic orbit for $g \in U_0$ in order to obtain properties for the periodic orbit, there is no essential difference between nonsingular endomorphisms and diffeomorphisms. Consequently, the basis of their results for diffeomorphisms such as Franks’ Lemma [7] (see also [14, Lemma II.2]) or Wen’s results [31, Lemmas 3.3 and 3.4] on periodic orbits (using the Franks’ Lemma along periodic orbits and diffeomorphisms. Consequently, the basis of their results for diffeomorphisms case, we can translate the properties on subbundles over periodic orbits given by [31, Lemmas 3.3 and 3.4] on periodic orbits (using the Franks’ Lemma along periodic orbits and creating homoclinic tangencies by local perturbations) still works, satisfying the weak hyperbolicity on subbundles over periodic orbits given by [31, Lemmas 3.3 and 3.4]. Then, applying Theorem A instead of Mañé’s Ergodic Closing Lemma in the same manner as in the diffeomorphisms case, we can translate the properties on subbundles over periodic orbits to those over the forward orbit of $\mu$-almost every $x$ for every $\mu \in M_f(g)$ to continue. However, for completeness and slight differences from the diffeomorphisms case, we give a proof of the lemma.

Proof of Lemma 3.5. As mentioned above, by the same argument to prove [31, Lemma 3.3 and 3.4] (see also [12, Proposition 4.1]), there exist a neighborhood $U_0$ of $f$ in $\text{NEnd}^1(M) \setminus \overline{HT}$, constants $\gamma_0 > 0$ and $m \in \mathbb{Z}^+$ such that for every $g \in U_0$ and every $p \in \text{Per}(g)$, letting $\hat{p} = (p_i)_{i \leq 0} \in S(g)^g_g$ be such that $\pi(\hat{p}) = p$ and $p_i \in O^+_g(p)$ for all $i \leq 0$, we have an $m$-dominated splitting

$$T_{S(g)}^{\gamma_0}(O^+_g(\hat{p})) = E^1_g \oplus E^2_g \oplus E^3_g$$

with $\dim E^2_g \leq 1$ whose nonzero vectors in $E^2_g$ have Lyapunov exponents in $[-\gamma_0, \gamma_0]$, satisfying that if $p$ has period $\tau \geq m$ then

$$\frac{1}{[\tau/m]} \sum_{j=0}^{[\tau/m]-1} \log \|Dg^m\|E^1_g(g^m(p))\| < -\gamma_0$$

(3.14)

and

$$\frac{1}{[\tau/m]} \sum_{j=0}^{[\tau/m]-1} m(Dg^m\|E^3_g(g^m(\hat{p})))^{-1} < -\gamma_0.$$  (3.15)

Here the angles between any two distinct subbundles of (3.13) are uniformly bounded away from zero by its $m$-domination (see [1, page 288] for elementary properties of dominated splittings).

Consider $\hat{x} = (x_i)_{i \leq 0} \in S(g)^g$ with $\pi(\hat{x}) = x \in \Sigma^e_+(g)$ satisfying

$$\mu = \lim_{n \to +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \delta_{\hat{p}(x)}$$
for some $\mu \in M_\epsilon(g)$. Then apply Theorem A to find $j_0 \geq 0$, $\tilde{g} \in \mathcal{U}_0$ arbitrarily close to $g$, and a periodic point $\tilde{p} \in \text{Per}(\tilde{g})$ with period $\tilde{\tau}$ such that $d((g^{j_0} + j_0)(x), \tilde{g}^j(\tilde{p}))$ is arbitrarily small for all $0 \leq j \leq \tilde{\tau}$, where $j_0$ does not depend on the choice of $\tilde{g}$ and $\tilde{p}$ by the definition of $\Sigma^+_\mu(g)$. By the Poincaré Recurrence Theorem, we may suppose $x \in \Sigma^+_\mu(g)$ is recurrent and therefore the periodic orbit $\mathcal{O}^+_\mu(\tilde{p})$ can be arbitrarily close to $\text{supp}(\mu)$ in the Hausdorff metric. Take sequences $\{g_n\}_{n \geq 1}$ and $\{p_n\}_{n \geq 1}$ of such $\tilde{g}$ and $\tilde{p} \in \text{Per}(\tilde{g})$ so that $g_n \to g$ in the $C^1$ topology and $(1/\tau_n) \sum_{j=0}^{\tau_n - 1} \delta_{g^j(p_n)} \to \mu$ in the weak* topology, where $\tau_n$ is the $g_n$-period of $p_n$. By continuity and accumulation, the sequence of $m$-dominated splittings over $\mathcal{O}^+_\mu(\tilde{p}_n)$, $n \geq 1$, given by (3.13) attaches to $\text{supp}(\mu)^g$ and hence to $S(g)^g$ an $m$-dominated splitting

$$T_{S(g)^g} = E^1_g \oplus E^2_g \oplus E^3_g$$

by ranging over $x \in \Sigma^+_\mu(g)$.

Now if (3.14) and (3.15) hold for $p \in \text{Per}(g)$ with period $\tau \geq m$, then replacing $m$ by some large multiple $m_0$ of $m$ if necessary, we see that the property of this lemma holds as $m_0$, $\rho_0 = \gamma_0$ and $\Lambda_0(g) = \Sigma^+_\mu(g)$.

Next, in the case where $\mu$ is supported on infinitely many points, we consider the above application of Theorem A through (3.16) similarly to the diffeomorphisms case as in [14, page 523]. Perturb $g_n$ with large $n$ slightly to $h_n$ by using Franks’ Lemma so that $\mathcal{O}^+_\mu(p_n) = \mathcal{O}^+_{h_n}(p_n)$ and $h_n \to g$ for which we have an $m$-dominated splitting

$$T_{S(h_n)^{\mu_0}(\tilde{p}_n)} = \tilde{E}^1_n \oplus \tilde{E}^2_n \oplus \tilde{E}^3_n$$

with $\dim \tilde{E}^2_n \leq 1$, satisfying (3.13), (3.14) and (3.15) with $g$ and $E^3_g$ replaced by $h_n$ and $E^3_n$, $i \in \{1, 2, 3\}$, and such that the properties of the derivatives on $\tilde{E}^1_n$ and $\tilde{E}^3_n$ over $\mathcal{O}^+_{h_n}(\tilde{p}_n)$ are isometrically translated to those on $E^1_g$ and $E^3_g$ over the string $(\tilde{g}^{j_0}(\tilde{x}), \tilde{g}^{\tau_n + j_0}(\tilde{x}); \tilde{g})$, respectively. Here the translation is achieved through isomorphisms $A_{n,j} : T_{h_n(p_n)}M \to T_{g^{j_0+n}(x)}M$ such that $E^1_{h_n}|\{h_n(\tilde{p}_n)\} = A^{-1}_{n,j}(E^1_g|\{\tilde{g}^{j_0}(\tilde{x})\})$, the restrictions to $E^1_{h_n}|\{h_n(\tilde{p}_n)\}$ are isometries for all $0 \leq j \leq \tau_n$ and

$$A^{-1}_{n,0}(E^3_g|\{\tilde{g}^{j_0}(\tilde{x})\}) = A^{-1}_{n,\tau_n}(E^3_g|\{\tilde{g}^{\tau_n + j_0}(\tilde{x})\}).$$

Since $j_0 \geq 0$ does not depend on the choice of $h_n$, and $\rho_n$, $n \geq 1$, when $\tau_n \to +\infty$ we have

$$\liminf_{\ell \to +\infty} \frac{1}{\ell} \sum_{j=0}^{\ell-1} \log \|Dg^m|E^1_g(g^{mj + j_0}(x))\| < -\gamma_0$$

and

$$\liminf_{\ell \to +\infty} \frac{1}{\ell} \sum_{j=0}^{\ell-1} \log \text{m}(D_{x_0}g^m|E^3_g(g^{mj + j_0}(\tilde{x})))^{-1} < -\gamma_0.$$  

By the ergodicity of $\mu$, the limit infimum of (3.17) can be changed to the limit, which together with (3.18) gives the required property in this case for some large multiple $m_0$ of $j_0(m + 1)m$, $\rho_0 = \gamma_0$ and $\Lambda_0(g) = \Sigma^+_\mu(g)$.

Finally, let us consider the case where $p \in \text{Per}(g)$ has period $\tau < m$. According to (3.13), the Lyapunov exponents for $g \in \mathcal{U}_0$ of nonzero vectors of $E^1_g$, $E^2_g$ and $E^3_g$ are in $(-\infty, -\gamma_0)$, $[-\gamma_0, \gamma_0]$ and $(\gamma_0, +\infty)$, respectively. Denote by $\text{Per}_{g}(g)$ the set of $p \in \text{Per}(g)$ with the minimum period $\ell$ and let $\hat{p} = (p_{i})_{i \leq 0} \in \text{Per}_{m}(g)^g$ be such that $\pi(\hat{p}) = p$ and $p_i \in \mathcal{O}^+_g(p)$ for all $i \leq 0$. By the uniformity of the constant $\gamma_0 > 0$
and angles between any two distinct subbundles of (3.16) for $g \in \mathcal{U}_0$, we cannot find $h$ arbitrarily $C^1$ close to $g$ with $\mathcal{O}_g^h(p) = \mathcal{O}_g(p)$ and $Dh^{-1}(E_{g}^2([\hat{p}])) = Dg^{-1}(E_{g}^2([\hat{p}]))$ such that $Dh^{-1}([E_{g}^2([\hat{p}])) : E_{g}^2([\hat{p}]) \rightarrow E_{g}^2([\hat{p}])$ or $Dh^{-1}([E_{g}^2([\hat{p}])) : E_{g}^2([\hat{p}]) \rightarrow E_{g}^2([\hat{p}])$ is a nonhyperbolic isomorphism. Then, define uniformly hyperbolic families $\{\xi_{i}^{(p,g,i)} : g \in \mathcal{U}_0, p \in \text{Per}_m(g)\}, i \in \{1,3\}$, of periodic sequence of isomorphisms $\xi_{i}^{(p,g,i)}$, $i \in \mathbb{Z}$, by:

$$\xi_{i}^{(p,g,i)} = \begin{cases} D_{g^i}(p)g[E_{g}^2([\hat{p}])], & \text{if } i \geq 1; \\ D_{p^i}g[E_{g}^2([p_{i+j}])], & \text{if } i \leq 0. \end{cases}$$

Now apply [14, Lemma II.3, (c)] to obtain that for every $\{\xi_{i}^{(p,g,i)} : g \in \mathcal{U}_0, p \in \text{Per}_m(g)\}, i \in \{1,3\}$, there exists $m_0 \in \mathbb{Z}^+$ such that

$$\limsup_{\ell \to +\infty} \frac{1}{\ell} \sum_{j=0}^{\ell-1} \log \left\| \prod_{i=0}^{m_0-1} \xi_{i+j}^{(p,g,1)} \right\| < 0 \quad (3.19)$$

and

$$\limsup_{\ell \to +\infty} \frac{1}{\ell} \sum_{j=0}^{\ell-1} \log m\left( \prod_{i=0}^{m_0-1} \xi_{i+j}^{(p,g,3)} \right)^{-1} < 0. \quad (3.20)$$

Here, in the proof of [14, Lemma II.3, (c)], setting

$$C = \sup \{\|\xi_{i}^{(p,g,1)}\| : g \in \mathcal{U}_0, p \in \text{Per}_m(g), i \in \mathbb{Z}\},$$

inequality (3.19) is proved from that for all $g \in \mathcal{U}_0, p \in \text{Per}_m(g)$ and $j \in \mathbb{Z}$ we have

$$\left\| \prod_{i=0}^{m_0-1} \xi_{i+j}^{(p,g,1)} \right\| \leq C^{m_0} (1/2)^{\lfloor m/m_0 \rfloor} \leq 1/2$$

for some positive integers $m_1$ and $\overline{m}$ (that may be any integer much larger than $m_1$) chosen independent of $\xi_{i}^{(p,g,1)}$ in the family $\{\xi_{i}^{(p,g,1)} : g \in \mathcal{U}_0, p \in \text{Per}_m(g)\}$ (see the proof of [14, Lemma II.7 (b)]). The proof of (3.20) is similar, so we may also assume that

$$m\left( \prod_{i=0}^{m_0-1} \xi_{i+j}^{(p,g,3)} \right)^{-1} \leq 1/2.$$

Then, choosing as $m_0$ a common multiple of $\overline{m}$ and $m$, we can find $\rho_0 > 0$ less than $\gamma_0$ such that inequalities (3.19) and (3.20) are actually written as:

$$\limsup_{\ell \to +\infty} \frac{1}{\ell} \sum_{j=0}^{\ell-1} \log \|Dg^{m_0}E_g^1(g^{m_0}(p))\| < -\rho_0$$

and

$$\limsup_{\ell \to +\infty} \frac{1}{\ell} \sum_{j=0}^{\ell-1} \log \left( D_{p^i}g^{m_0}E_g^3(g^{m_0}(\hat{p})) \right)^{-1} < -\rho_0.$$

Thus, if $m_0$ has been chosen as some large multiple of $m$, the required property holds including the case where $p \in \text{Per}_m(g)$ as $\Lambda_0(g) = \Sigma_+^m(g)$. \hfill \Box

Applying Lemma 3.1 to $a = 0$ with Proposition 3.6, we see that if $f \in \text{NEnd}_1^1(M) \setminus \mathcal{HT}$, there exist a neighborhood $\mathcal{U}$ of $f$ and $\varepsilon > 0$ such that every $g \in \mathcal{U}$ is $\varepsilon$-entropy expansive to obtain Theorem B. Then (2.3) implies the following theorem.
There exist $i_0 \leq z,y$ we can apply perturbation (II) to $(\text{II})$ for every $q \in S(f)$ satisfying $\|q\| \leq r_1/2$, there exists $0 < \varepsilon(q,J) < r_1/2$ such that $f^j(B_{3\varepsilon(q,J)}(q))$ is a diffeomorphism onto its image for all $0 \leq j \leq J$ and if max$\{||y - q||, ||z - q||\} \leq \varepsilon(q,J)$, then there exists $g \in \mathcal{U}$ such that $g^j(y) = f^j(z)$

and $g(w) = f(w)$

when

$$w \notin \bigcup_{j=1}^{j-1} f^j\left(\text{Int } \tilde{B}_{\delta r}\left(\frac{y+z}{2}\right)\right),$$

where $r = 2^{-1}||y-z||$ and $\tilde{B}_t(v) = \{x : ||x-v|| \leq t\}$.

Restricting the observation of [9, (2.2)] over $M$ to $S(f)$, we can take $\{p_1, \ldots, p_k\}$ such that

$$S(f) \subset \bigcup_{i=1}^{k} \text{Int } B_{r_0(p_i)/2}(p_i) \subset \bigcup_{i=1}^{k} U_0(p_i)$$

with one perturbation framework in each $B_{r_0(p_i)}(p_i)$ given as (I) and (II).

Given a neighborhood $\mathcal{U}$ of $f \in \text{End}^1(M)$ and $\varepsilon > 0$, we say that a pair of points $(z,y)$ with $y = f^\ell(z)$ for some $\ell \in \mathbb{Z}^+$ is a $(\mathcal{U}, \varepsilon)$-strongly closable pair of $f$ if $y = z$ or $y \neq z$ satisfies properties (E-a), (E-b) and (E-c) below. Then, since

$$\tilde{B}_{3\varepsilon(q,J)}(q),$$

we can apply perturbation (II) to $(z,y)$ in order to obtain a periodic orbit $\{g^j(y) : 0 \leq j \leq \ell - 1\}$ with period $\ell$ for some $q \in \mathcal{U}$.

(E-a) There exist $i \in \{1, \ldots, k\}$, $q \in B_{r_0(p_i)/2}(p_i) \setminus \text{Per}_{J_i}(f)$ and $\varepsilon_i(q,J_i) > 0$ such that max$\{||y - q||, ||z - q||\} \leq \varepsilon_i(q,J_i)$ and satisfies

$$f^j\left(\tilde{B}_{3\varepsilon_i(q,J_i)}(q)\right) \cap f^j\left(\tilde{B}_{3\varepsilon_i(q,J_i)}(q)\right) = \emptyset$$
For all $j \neq j'$, for $p = p_i$, $\delta = \delta_i$, $J = J_i$ (from which $\| \cdot \|_1$ is taken) and $\varepsilon(q, J) = \varepsilon_i(q, J_i)$.

(E-b) For all $J_1 \leq j \leq \ell$,

$$f^j(z) \notin \bigcup_{j=1}^{J_1-1} f^j\left(\text{Int} \, B_{\delta,r}\left(\frac{y+z}{2}\right)\right)$$

with $r = 2^{-1}\|y-z\|_1$.

(E-c) For the same $r$ as above,

$$\max_{0 \leq j \leq J_i} \text{diam} f^j\left(\overline{B}_{\delta,r}\left(\frac{y+z}{2}\right)\right) \leq \varepsilon.$$

Remarks. (1) Similarly to the diffeomorphisms case, the norm $\| \cdot \|_1$ above actually depends on the configuration of $(z, y)$ in a uniform way (see [14] or [22]). For simplicity, we omit the dependence here.

(2) Since $f f^j(\text{Int} \, B_{\delta,r}(\frac{y+z}{2}))$ is a diffeomorphism onto its image for all $0 \leq j \leq J_i$, there is no difference in the closing perturbation between diffeomorphisms and endomorphisms as long as (E-b) is satisfied.

(3) Call (D-a), (D-b) and (D-c), respectively, the properties (a), (b) and (c) in [9, page 784] for diffeomorphisms. Then (D-a) and (D-c) are the same as (E-a) and (E-c), and the only difference from the diffeomorphisms case is (E-b). In fact, for $f \in \text{End}^1(M)$ it can happen that $\{f^j(z) : J_i \leq j \leq \ell\}$ intersects $\bigcup_{j=1}^{J_i-1} f^j(\text{Int} \, B_{\delta,r}(\frac{y+z}{2}))$ without hitting $\text{Int} \, B_{\delta,r}(\frac{y+z}{2})$. This interference is the new difficulty in the endomorphisms case. This motivates us to change (D-b) to (E-b) in order to avoid the interference.

Given $\ell \in \mathbb{Z}^+$, denote by $\Sigma^-(\mathcal{U}, \varepsilon, \ell)$ the set of $z$ by which $(z, y)$ becomes a $(\mathcal{U}, \varepsilon)$-strongly closable pair for some $y$ with $y = f^\ell(z)$. Then, define

$$\Sigma^-(\mathcal{U}, \varepsilon) = \bigcup_{\ell \geq 1} \Sigma^-(\mathcal{U}, \varepsilon, \ell).$$

Let

$$J = \max_{1 \leq j \leq k} J_j.$$

Define a Borel set $\Sigma_j(\mathcal{U}, \varepsilon)$ as the set of points $x \in M$ such that there exist $g \in \mathcal{U}$, $\ell \in \mathbb{Z}^+$ and $y \in \text{Per}_r(g)$ satisfying $g = f$ on $M \setminus B_{\ell}(f, x, J)$ and $d(f^j(x), g^j(y)) \leq \varepsilon$ for all $0 \leq j \leq \ell$, where $B_{\ell}(f, x, l)$ with $l \geq 1$ is the set of $w \in M$ such that $d(f^j(x), w) \leq \varepsilon$ for some $0 \leq j \leq l$. Note that

$$\Sigma^-(\mathcal{U}, \varepsilon) \subset \Sigma_j(\mathcal{U}, \varepsilon) \subset \Sigma_+(\mathcal{U}, \varepsilon).$$

By the Ergodic Decomposition Theorem, the proof of Theorem A is reduced to proving that given a neighborhood $\mathcal{U}$ of $f$,

$$\mu(\Sigma^e_+(\mathcal{U}, f)) = 1$$

for every $\mu \in \mathcal{M}_c(f)$. For the proof, given $x$ in a total probability set we need to have some $m \geq 0$ such that $f^m(x) \in \Sigma_+(\mathcal{U}, \varepsilon)$ for every $\varepsilon > 0$. The nontrivial case is when $\mu$ is supported on infinitely many points and then finding $f^m(x)$ as a $(\mathcal{U}, \varepsilon)$-strongly closable point for every $\varepsilon > 0$ suffices to prove (4.1). As the previous step to find $(\mathcal{U}, \varepsilon)$-strongly closable pairs, we define pairs satisfying weaker properties. Given a neighborhood $\mathcal{U}$ of $f$ and $\varepsilon > 0$, we say that a pair of points $(z, y)$ with $y = f^\ell(z)$ for some $\ell \in \mathbb{Z}^+$ is a $(\mathcal{U}, \varepsilon)$-quasi-closable pair of $f$ if $y = z$ or $y \neq z$
satisfies (E-c) in addition to the following property (D-b), one of the properties required for \((U, \varepsilon, \rho, m, \ell, w, B, f)\)-strongly closable pairs \((y, z)\) for diffeomorphisms in [9]:

(D-b) For all \(0 < j < \ell\),

\[
f^j(z) \notin \text{Int} B_{8, r} \left( \frac{y + z}{2} \right)
\]

with \(r = 2^{-1} \|y - z\|_1\).

In particular, when \(f \in \text{End}^1(M)\) is a diffeomorphism, adding property (E-a) to a \((U, \varepsilon, \rho, m, \ell, w, B, f)\)-quasi-closable pair \((z, y)\) makes it \((U, \varepsilon, \rho, m, \ell, w, B, f)\)-strongly closable. But, as mentioned in (3) of the Remarks above, it is not sufficient to be \((U, \varepsilon, \rho, m, \ell, w, B, f)\)-quasi-closable for endomorphisms.

Denote by \(\Sigma_{ij}(U, \varepsilon)\) the set of \(z\) for which \((z, y)\) becomes a \((U, \varepsilon, \rho, m, \ell, w, B, f)\)-quasi-closable pair for some \(y\). Then, by the same argument as in the proof of Mañé’s Ergodic Closing Lemma ([14], see also [9]), it is easy to see that \(\Sigma_{ij}(U, \varepsilon)\) is a Borel set and

\[
\mu(\Sigma_{ij}(U, \varepsilon)) = 1
\]

for every \(\mu \in \mathcal{M}_c(f)\).

The following selection lemma is the fundamental procedure for the closing, which goes back to Pugh’s Closing Lemma [21]. Mañé investigated Pugh’s pointwise selection lemma from a global viewpoint in his proof of the Ergodic Closing Lemma [14], dividing the ambient manifold to countable pieces around which the specified smallness and the shapes of boxes for perturbation (II) are available. In [9], it was observed that the finite choice of shapes of the boxes is enough as long as the sufficient smallness is guaranteed. The finiteness is used in the following selection lemma in order to give \(\rho > 1\) independent of the choice of \(q\) around which pairs are selected. However, the selection itself with the norm \(\| \cdot \|_1\) is the same as one used in the proof of the Ergodic Closing Lemma. So, we omit the proof (see [14, Lemma 1.4]).

**Lemma 4.1 (Selection Lemma).** Let \(f \in \text{End}^1(M)\) satisfy \(S(f) \cap C(f) = \emptyset\). Given a neighborhood \(U\) of \(f\) and \(\varepsilon > 0\), there exist \(\rho > 1\) and \(0 < r_0 = r_0(\rho) < \min\{r_1(p_i)/2 : 1 \leq i \leq k\}\) such that if \(q \in B_{r_1(p_i)}/2(p_i)\) and \(\{x, f^m(x)\} \subset B_r(q)\) with some \(m > 0\) and \(0 < r < r_0\), then there are \(0 \leq m_1 < m_2 \leq m\) such that \(\{f^m_1(x), f^m_2(x)\} \subset B_{mp}(q) \subset B_{r_1(p_i)}(p_i)\) and the pair \((f^m_1(x), f^m_2(x))\) satisfying (E-c) is \((U, \varepsilon, \rho, m, \ell, w, B, f)\)-quasi-closable.

**Remark.** If \(r > 0\) is much smaller than \(\min\{\varepsilon, \min_{1 \leq i \leq k} r_1(p_i)/2\}\) then the selection to obtain \((f^m_1(x), f^m_2(x))\) as above around \(q\) by the norm \(\| \cdot \|_1\) always works to get properties (D-b) and (E-c). However, we don’t know if \(r > 0\) is so small to satisfy (E-c). This is the reason for assuming (E-c) in the conclusion. On the other hand, to get also property (E-a) we need smallness of pairs determined by \(\varepsilon_i(q, J_i)\), which is not assumed in Lemma 4.1.

Let us give the notation and properties involved in the proof of the Ergodic Closing Lemma [14] (see also [9]). For a neighborhood \(U\) of \(f\) in \(\text{End}^1(M)\), \(\varepsilon > 0\), \(r > 0\) and \(\rho > 1\), define a Borel set \(\Sigma(U, \varepsilon, r, \rho, m, \ell)\) as the set of points \(x \in M\) such that if \(w \in B_r(x)\) and \(f^m(w) \in B_r(x)\) for some \(0 < \tilde{r} \leq r\) and \(m \in \mathbb{Z}^+\), then there exist \(0 \leq m_1 < m_2 \leq m\) such that \((f^m_1(w), f^m_2(w))\) is a \((U, \varepsilon, \rho, m, \ell, w, B)\)-quasi-closable pair with \(f^m_1(w), f^m_2(w) \in B_{mp}(x)\). Note that this property is weaker than the property required for the set written by the same notation in [14] even for diffeomorphisms. If \(r_n > 0, n \geq 0\), is a monotone decreasing sequence converging
to 0, then \( \Sigma(\mathcal{U}, \varepsilon, r_n, \rho) \), \( n \geq 0 \), is an increasing sequence and it is known ([14, 9]) that
\[
S(f) = \bigcup_{n \geq 0} \Sigma(\mathcal{U}, \varepsilon, r_n, \rho).
\] (4.3)

The fact that \( \rho > 1 \) can be fixed was observed in [9] (while Mañé used a sequence \( \rho_m \to +\infty \)). This property corresponds to the uniformity of shapes of perturbation boxes (see [14]) from which \( J \in \mathbb{Z}^+ \) is fixed after \( \mathcal{U} \) was fixed first. The advantage is that smallness \( r_n \) for \((\mathcal{U}, \varepsilon)\)-quasi-closable pairs with property (E-b) to be \((\mathcal{U}, \varepsilon)\)-strongly closable can be uniformly chosen outside a neighborhood of \( \text{Per}_J(f) \).

For the measure-theoretical argument in the proof of the Ergodic Closing Lemma, Mañé introduced a sequence of partition of \( M \) as follows. The ambient manifold \( M \) is isometrically embedded in \( T^s = S^1 \times \cdots \times S^1 \) and a cube \( A \) as an atom of the partition of \( T^s \) is defined by \( A = I_1 \times \cdots \times I_s \) where \( I_i \), \( 1 \leq i \leq s \), are intervals in \( S^1 \) with equal lengths. If \( p_i \in I_i \) is the middle point of \( I_i \), then \((p_1, \ldots, p_s)\) is called the center of \( A \). For each \( k \in \mathbb{Z}^+ \), let \( \mathcal{P}_1^{(k)} \leq \mathcal{P}_2^{(k)} \leq \cdots \) be a sequence of partitions, where \( \mathcal{P}_j^{(k)} \) is the partition whose each atom \( A \) have the side \( 2\pi/k^j \) (the length of the interval \( I_i \)). Then, denote by \( \hat{A} \) and \( \hat{\hat{A}} \) the cubes with the same center as \( A \)'s with the sides \( 2\pi/k^{j-1} \) and \( 6\pi/k^{j-1} \), respectively. If \( x \in T^s \), we denote by \( \mathcal{P}_j^{(k)}(x) \) the atom of the partition containing \( x \). Then the following quite general property holds.

**Lemma 4.2** ([14, Lemma I.5]). Let \( \mu \in \mathcal{M}(T^s) \), \( \delta > 0 \) and an odd integer \( k \) be given. Then for every \( j \geq 1 \) the following inequalities hold:
\[
\mu\left( \{ x \in T^s : \mu(\mathcal{P}_j^{(k)}(x)) \geq \delta \mu(\mathcal{P}_j^{(k)}(x)) \} \right) \geq 1 - \delta k^s
\]
and
\[
\mu\left( \{ x \in T^s : \mu(\mathcal{P}_j^{(k)}(x)) \geq \delta \mu(\mathcal{P}_j^{(k)}(x)) \} \right) \geq 1 - \delta 3^s k^s.
\]

In order to consider \( \mu \in \mathcal{M}(M) \) with the partition of \( T^s \), we extend \( \mu \) to a measure on \( T^s \) by defining \( \mu(B) = \mu(B \cap M) \) for every Borel set \( B \) of \( T^s \).

Now let \( \mu \in \mathcal{M}_c(f) \) be supported on infinitely many points. Given \( \rho > 1 \) and \( n \geq 0 \), there is \( j(n) \geq 1 \) such that for all \( j \geq j(n) \) take \( 0 < r \leq r_n \) and an odd integer \( k \geq 1 \) satisfying
\[
\mathcal{P}_j^{(k)}(x) \subset B_r(x) \quad \text{and} \quad \mathcal{P}_j^{(k)}(x) \supset B_{\rho r}(x),
\] (4.4)
where \( B_r(w) \) is the closed ball in \( T^s \) with radius \( t \) and center \( w \). Note that if \( w \in M \) then \( B_r(w) \cap M = B_r(w) \). Without loss of generality, we may assume that
\[
\mu(\partial \mathcal{P}_j^{(k)}(x)) = \mu(\partial \mathcal{P}_j^{(k)}(x)) = \mu(\partial \mathcal{P}_j^{(k)}(x)) = 0
\]
for all \( k, j \) and \( x \in M \). The following lemma corresponds to [14, Lemma I.6] and its proof is the same as that of [14, Lemma I.6], giving only a sketch of the proof for later arguments.

**Lemma 4.3.** If \( x \in \Sigma(\mathcal{U}, \varepsilon, r_n, \rho) \), \( j \geq j(n) \) and \( \mu(\mathcal{P}_j^{(k)}(x)) \geq \delta \mu(\mathcal{P}_j^{(k)}(x)) \) with \( k \) satisfying (4.4), then
\[
\mu(\mathcal{P}_j^{(k)}(x) \cap \Sigma(f, \varepsilon)) \geq \delta \mu(\mathcal{P}_j^{(k)}(x)).
\]
Sketch of Proof. Since $\mu$ is ergodic, there is $x_0 \in M$ such that the hitting frequency in $P_j^{(k)}(x) \cap \hat{\Sigma}_j(U, \varepsilon)$ and $P_j^{(k)}(x)$ of $\{f(x_0), \ldots, f^{\ell}(x_0)\}$ over $\ell$ converge to $\mu$-measures of $P_j^{(k)}(x) \cap \hat{\Sigma}_j(U, \varepsilon)$ and $P_j^{(k)}(x)$ as $\ell \to +\infty$, respectively. If there are no consecutive points in the forward orbit of $x_0$ hitting $P_j^{(k)}(x)$, we can choose a $(U, \varepsilon)$-quasi-closable pair in $\hat{P}_j^{(k)}(x)$ from the string consisting of the consecutive points in $P_j^{(k)}(x)$ and points in between them (with respect to the forward orbital order). This means that the number of points in $\{f(x_0), \ldots, f^{\ell}(x_0)\}$ hitting $P_j^{(k)}(x)$ is the same as the number of their hitting $\hat{P}_j^{(k)}(x) \cap \hat{\Sigma}_j(U, \varepsilon)$. Consequently, it holds that either $\hat{P}_j^{(k)}(x)$ or $P_j^{(k)}(x)$ is $\delta$-compatible to that of $P_j^{(k)}(x)$ or $\hat{P}_j^{(k)}(x)$, the hitting frequency of the forward orbit of $x_0$ in $P_j^{(k)}(x) \cap \hat{\Sigma}_j(U, \varepsilon)$ is $\delta$-compatible to that of $x_0$ in $\hat{P}_j^{(k)}(x)$, obtaining the required inequality.

Now we need to change some of $(U, \varepsilon)$-quasi-closable pairs to $(U, \varepsilon)$-strongly closable ones. The following lemma is the first step.

Lemma 4.4. Let $(z, y)$ be a $(U, \varepsilon)$-quasi-closable pair with $z \in \hat{\Sigma}_j(U, \varepsilon)$ and $y = f^i(z)$ for some $\ell > 0$. If $0 < \varepsilon < \tau_0$, then there exist $0 \leq s_1 < s_2 \leq \ell$ with $\{f^{s_1}(z), f^{s_2}(z)\} \subset B_{r_1(p)}(p_1)$ for some $1 \leq i \leq k$ such that if the pair $(f^{s_1}(z), f^{s_2}(z))$ is $(U, \varepsilon)$-quasi-closable then it satisfies properties (E-b) and (E-c). Consequently, it holds that either $f^{s_1}(z) \in \Sigma^{-}(U, \varepsilon)$ or property (E-a) is not satisfied for the $(U, \varepsilon)$-quasi-closable pair $(f^{s_1}(z), f^{s_2}(z))$.

Proof. If the given $(U, \varepsilon)$-quasi-closable pair $(z, y)$ satisfies property (E-b), we have $z \in \Sigma^{-}(U, \varepsilon)$ or property (E-a) is not satisfied for $(z, y)$. Then $s_1 = 0$ and $s_2 = \ell$ is the required one. So let us suppose $(z, y)$ does not satisfy property (E-b). Then there exist $1 \leq i_0 \leq k$, $0 < j_0 < J_{i_0}$ and $J_{i_0} \leq t_0 \leq \ell$ such that

$$f^{t_0}(z) \in f^{j_0}(\text{Int} \hat{B}_{\delta_0R}\left(\frac{y + z}{2}\right))$$

with $r = 2^{-1}\|y - z\|_1$. By property (E-c) for $(z, y)$, there is $1 \leq i \leq k$ such that

$$f^{j_0}(z) \in B_{r_1(p_1)}/2(p_1) \quad \text{and} \quad d(f^{j_0}(z), f^{j_0}(z)) = \varepsilon.$$

Now, if $0 < \varepsilon < \tilde{r}_0$ then, by Lemma 4.1 with $x = q = f^{j_0}(z)$, $m = t_0 - j_0$ and $r = \varepsilon$, there exist $t_0 \leq t_1 < t_2 \leq \ell$ such that, setting $(z', y') = (f^{t_1}(z), f^{t_2}(z))$, if $(z', y')$ satisfies property (E-c) then we have a pair $(z', y') \subset B_{r_1(p_1)}(p_1)$ with properties (D-b) and (E-c), becoming $(U, \varepsilon)$-quasi-closable.

Remark. The hypothesis that $(f^{s_1}(z), f^{s_2}(z))$ is $(U, \varepsilon)$-quasi-closable seems superfluous at first sight. However, because of the selection process using the norm $\| \cdot \|_1$, the size of the pair $\|f^{s_1}(z) - f^{s_2}(z)\|_1$ is not clear although pairs necessary to us are only small ones (see the Remark below Lemma 4.1). We will see later that large pairs can be neglected.
The next step for \((\mathcal{U}, \varepsilon)\)-quasi-closable pairs to be \((\mathcal{U}, \varepsilon)\)-strongly closable is obtaining property (E-a). Let \(x_0 \notin \text{Per}(f)\) be \(\mu\)-almost every point for \(\mu \in \mathcal{M}_r(f)\) and let \(\ell_i, i \geq 1\), be a monotone increasing sequence such that \(\ell_i : i \geq 1\) is the set of \(j \geq 1\) for which \(f^j(x_0) \in \Sigma_J(\mathcal{U}, \varepsilon)\) with \(\varepsilon < \xi_0\). Then there are \(k_i > \ell_i, i \geq 1\), such that \((f^{k_i}(x_0), f^{k_i}(x_0))\) is \((\mathcal{U}, \varepsilon)\)-quasi-closable. By Lemma 4.4, there exist pairs \((f^{k_i}(x_0), f^{k_i}(x_0))\) with \(\ell_i \leq t_i < u_i \leq k_i, i \geq 1\), such that if it is \((\mathcal{U}, \varepsilon)\)-quasi-closable and \(0 < \varepsilon < \xi_0\) then either \(f^{k_i}(x_0) \in \Sigma^- (\mathcal{U}, \varepsilon)\) or \((f^{k_i}(x_0), f^{k_i}(x_0))\) with properties (E-b) and (E-c) does not satisfy property (E-a).

We consider two cases according to the positions of \(f^{k_i}(x_0), i \geq 1\). To simplify the notation, put \((z', y') = (f^{k_i}(x_0), f^{k_i}(x_0))\), which will be called a pair chosen from \((f^{k_i}(x_0), f^{k_i}(x_0))\). For every \(x_i \in \text{Per}_J(f)\) with period \(\ell \leq J\), we can find a neighborhood \(V(x_i)\) of \(x_i\) such that \(f|f^j(V(x_i))\) is a diffeomorphism onto its image for all \(0 \leq j \leq J\), and if \(x \in V(x_i)\) then

\[
d(f^j(x), f^j(x)) \leq \varepsilon
\]

for all \(0 \leq j \leq \ell\). Since \(\text{Per}_J(f)\) is compact, we can take a finite subcover \(\{V(x_i) : 1 \leq l \leq \tau\}\) of \(\{V(x_i) : x_i \in \text{Per}_J(f)\}\) and let the union be \(V_J\), i.e.,

\[
\text{Per}_J(f) \subset \bigcup_{l=1}^{\tau} V(x_i) = V_J.
\]

Then, if \(z' \in V_J\), one can find \(x_i \in \text{Per}_J(f)\) with period \(\ell \leq J\) satisfying \(d(f^j(z'), f^j(x_i)) \leq \varepsilon\) for all \(0 \leq j \leq \ell\). By the presence of \(\{f^j(z') : 0 \leq j \leq \ell\}\), we see that

\[
z' = f^{k_i}(x_0) \in \Sigma_J(\mathcal{U}, \varepsilon).
\]

(4.5)

On the other hand, for every \(q \in S(f) \setminus V_J\), there is a neighborhood \(U(q)\) such that \(f|f^j(U(q))\) is a diffeomorphism onto its image for all \(0 \leq j \leq J\). Then we can take \(\varepsilon_i(q, J_i) > 0\) given in (E-a) with \(q \in B_{r_i(p_i)}(p_i)\) for some \(1 \leq i \leq k\) so small that \(\text{Int} \bar{B}_{\varepsilon_i(q, J_i)(q)} \subset U(q)\). Then set

\[
W(q) = \text{Int} \bar{B}_{\varepsilon_i(q, J_i)(q)}(q).
\]

Given \(0 < \varepsilon_0 \leq \varepsilon\), there is \(0 < \delta_0 = \delta_0(\varepsilon_0) \leq \varepsilon_i(q, J_i)/2\) such that if a pair \((z', y')\) with properties (E-b) and (E-c) satisfies \(\|z' - y'\|_1 < \delta_0\) and \(z' \in W(q)\), then the pair \((z', y')\) also satisfies property (E-a), becoming a \((\mathcal{U}, \varepsilon_0)\)-strongly closable pair. Since \(S(f) \setminus V_J\) is compact, we can take a finite subcover \(\{W(q_i) : 1 \leq l \leq t\}\) of \(\{W(q) : q \in S(f) \setminus V_J\}\) and let the union be \(V_J\), i.e.,

\[
S(f) \setminus V_J \subset \bigcup_{l=1}^{t} W(q_i) = W_J,
\]

and let

\[
\eta = \min_{1 \leq l \leq t} \varepsilon_i(q_i, J_i(q_i))/2.
\]

Then, if \(0 < \delta_0 = \delta_0(\varepsilon_0) \leq \eta\) and a pair \((z', y')\) with properties (E-b) and (E-c) satisfies \(\|z' - y'\|_1 < \delta_0\) and \(z' \in W_J\), property (E-a) is also satisfied to have \((z', y')\) as a \((\mathcal{U}, \varepsilon_0)\)-strongly closable pair with \(z' \in \Sigma^- (\mathcal{U}, \varepsilon_0)\). From this and Lemma 4.4 it follows that when we let \(\{(f^{k_i}(x_0), f^{k_i}(x_0)) : 1 \leq i \leq \xi_0(\ell)\}\) be the set of all \((\mathcal{U}, \varepsilon)\)-quasi-closable pairs with \(\varepsilon < \xi_0\) appeared in \(\{(f^{k_i}(x_0), f^{k_i}(x_0)) : 1 \leq i \leq \ell\}\) for large \(\ell\), there exist \(\ell_i \leq t_i < u_i \leq k_i \leq \ell\) such that either \(f^{k_i}(x_0) \notin W_J\) or if \(f^{k_i}(x_0) \in W_J\) then

\[
f^{k_i}(x_0) \in \Sigma_J(\mathcal{U}, \varepsilon_0).
\]

(4.6)
Define $\hat{\Lambda}_{\delta,n}$, for $\delta > 0$, as the set of points $x \in T^s$ such that for $k$ satisfying (4.4) the inequalities
\[
\mu(P_j^{(k)}(x)) \geq \delta \mu(\hat{P}_j^{(k)}(x))
\]
and
\[
\mu(P_j^{(k)}(x)) \geq \delta \mu(\hat{P}_j^{(k)}(x))
\]
hold for an infinite sequence $\nu(x)$ of values of $j \geq j(n)$. Then, by Lemma 4.2, we have an increasing sequence $\hat{\Lambda}_{1/m,n}$ with respect to $m \geq 1$ such that
\[
\mu(\hat{\Lambda}_{1/m,n}) \geq 1 - m^{-1}(k^s + 3^sk^s).
\]
Define $\Lambda_{\delta,n} = \Sigma(\mathcal{U}, \varepsilon, r_n, \rho) \cap \hat{\Lambda}_{\delta,n}$. Then from this inequality and (4.3) it follows that
\[
\bigcup_{n \geq 0} \bigcup_{m \geq 1} \Lambda_{1/m,n} = \bigcup_{n \geq 0} \Sigma(\mathcal{U}, \varepsilon, r_n, \rho) \cap S(f) \mod(0).
\]
Since $\Lambda_{1/m,n}$ is increasing in both $n \geq 0$ and $m \geq 1$, the proof of (4.1) is reduced to proving that for any $a > 0$ there exist arbitrarily large $m$ and $n$ such that
\[
\mu(\Sigma^c_+(\mathcal{U}, f)^c \cap \Lambda_{1/m,n}) \leq a. \tag{4.7}
\]
For the proof of this inequality, we need the following lemma corresponding to [14, Lemma I.7], whose proof is omitted, and it is the same as that of [14, Lemma I.7] where $\Sigma(\mathcal{U}, \varepsilon)^c$ is used instead of $\Sigma^c_+(\mathcal{U}, f)^c$. In fact, we can replace $\Sigma(\mathcal{U}, \varepsilon)^c$ by any Borel set in the proof of [14, Lemma I.7].

**Lemma 4.5.** Given $m, n \geq 0$ and a neighborhood $\mathcal{U}$ of $\Sigma^c_+(\mathcal{U}, f)^c \cap \Lambda_{1/m,n}$, there exist sequences $x_i \in \Sigma^c_+(\mathcal{U}, f)^c \cap \Lambda_{1/m,n}$, $j_i \in \nu(x_i)$ and $\hat{P}_j^{(k)}(x_i) \subset \mathcal{U}$, disjoint in mod(0) for $i = 1, 2, \ldots$, and such that
\[
\mu\left(\Sigma^c_+(\mathcal{U}, f)^c \cap \Lambda_{1/m,n} \setminus \bigcup_i \hat{P}_j^{(k)}(x_i)\right) = 0.
\]

If (4.7) does not hold, there exist $a_0 > 0$ and $m_0, n_0 \geq 1$ such that for any $m \geq m_0$ and $n \geq n_0$ we have
\[
\mu(\Sigma^c_+(\mathcal{U}, f)^c \cap \Lambda_{1/m,n}) > a_0.
\]
To lead a contradiction from this inequality, we will find a sufficient quantity of points in $\Sigma^c_+(\mathcal{U}, f)$ around $\Sigma^c_+(\mathcal{U}, f)^c \cap \Lambda_{1/m,n}$. By Lemma 4.5, for fixed such $m$ and $n$, we have an arbitrarily small neighborhood $\mathcal{U}$ of $\Sigma^c_+(\mathcal{U}, f)^c \cap \Lambda_{1/m,n}$ with $\mathcal{U} \subset V_J \cup W_J$, a positive integer $N \geq 1$ and cubes $\hat{P}_j^{(k)}(x_i) \subset \mathcal{U}$, $i = 1, 2, \ldots, N$ as in Lemma 4.5, satisfying
\[
\mu\left(\hat{S}_J(\mathcal{U}, \varepsilon) \cap \bigcup_{i=1}^{N} \hat{P}_j^{(k)}(x_i)\right) \leq a_0, \tag{4.8}
\]
where one can recall from (4.2) that $\mu(\hat{S}_J(\mathcal{U}, \varepsilon)) = 1$. Take a neighborhood $\mathcal{U}_N$ of $\Sigma^c_+(\mathcal{U}, f)^c \cap S(f)$ with $\mu(\partial\mathcal{U}_N) = 0$ and $\mathcal{U}_N \subset V_J \cup W_J$ such that
\[
\mu(\Sigma^c_+(\mathcal{U}, f) \cap \mathcal{U}_N) < \frac{a_0}{16mN} \left(1 - \frac{1}{2m}\right). \tag{4.9}
\]
If $\mathcal{U}$ has been taken small enough so that it contains a small amount of points in $\Sigma^c_+(\mathcal{U}, f)$ depending on $m$, we have
\[
\mu(\mathcal{U} \cap \mathcal{U}_N) < \frac{a_0}{4m}. \tag{4.10}
\]
To simplify the notation, we put $A_i = \mathcal{P}_j^{(k)}(x_i)$ and $\hat{A}_i = \mathcal{P}_j^{(k)}(x_i)$, $i = 1, 2, \ldots, N$, with $x_i \in A_{1/m,n}$. We need the following lemma in order to restrict the argument to $U_N$ and use (4.9).

**Lemma 4.6.** Let $\mathcal{I}_N = \{1 \leq i \leq N : \mu(A_i \cap U_N) < (2m)^{-1}\mu(\hat{A}_i \cap U_N)\}$. Then

$$\mu\left( \bigcup_{i \in \mathcal{I}_N} \hat{A}_i \cap U_N \right) < \frac{a_0}{2}.$$

**Proof.** Since $x_i \in A_{1/m,n}$, $i = 1, 2, \ldots, N$, we have

$$\mu(A_i \cap U_N) + \mu(A_i \cap U_N^c) = \mu(A_i) \geq m^{-1}\mu(\hat{A}_i) \geq m^{-1}\mu(\hat{A}_i \cap U_N)$$

for all $1 \leq i \leq N$. In particular, if $i \in \mathcal{I}_N$ then $\mu(A_i \cap U_N) < (2m)^{-1}\mu(\hat{A}_i \cap U_N)$, and therefore

$$\mu(A_i \cap U_N^c) \geq (2m)^{-1}\mu(\hat{A}_i \cap U_N).$$

Now if this lemma is false, we have

$$\mu\left( \bigcup_{i \in \mathcal{I}_N} \hat{A}_i \cap U_N \right) \geq \frac{a_0}{2},$$

and hence

$$\mu(U \cap U_N^c) \geq \mu\left( \bigcup_{i \in \mathcal{I}_N} A_i \cap U_N^c \right) = \sum_{i \in \mathcal{I}_N} \mu(A_i \cap U_N^c) \geq \sum_{i \in \mathcal{I}_N} \frac{1}{2m} \mu(\hat{A}_i \cap U_N) \geq \frac{1}{2m} \mu\left( \bigcup_{i \in \mathcal{I}_N} \hat{A}_i \cap U_N \right) \geq \frac{a_0}{4m},$$

contradicting (4.10). \(\square\)

Define $A_N = \bigcup_{i \notin \mathcal{I}_N} \hat{A}_i$. In the proof of Lemma 4.3, a $(U, \varepsilon)$-quasi-closable pair $(z, y)$ in $\hat{A}_i$ was chosen from any string consisting of consecutive points in $A_i$ and points in between them (with respect to the forward orbital order) of $\{f^j(x_0) : 1 \leq j \leq \ell\}$ with large $\ell \geq 1$. Then, from (4.5) and (4.6) we can find a pair $(z', y')$ chosen from $(z, y)$ such that $z' \in \Sigma_j(U, \varepsilon)$ if $z' \notin W_j$ (and therefore $z' \in V_j$), and given $0 < \varepsilon_0 < \varepsilon$, if $(z', y')$ is sufficiently small $(U, \varepsilon)$-quasi-closable pair with $z' \in W_j$ then $z' \in \Sigma_j(U, \varepsilon_0)$, where $(z, y)$ and $(z', y')$ correspond to $(f^{\ell_i}(x_0), f^{\ell_i}(x_0))$ and $(f^{\ell_i}(x_0), f^{\ell_i}(x_0))$ for some $1 \leq i \leq \zeta_0(\ell)$, respectively, given above (4.6). Here we need to know that the cardinality of the set of such pairs $(z', y')$ is comparable to the number of $(U, \varepsilon)$-quasi-closable pairs $(z, y)$ in $\{f^j(x_0) : 1 \leq j \leq \ell\} \cap A_N \cap U_N$ from which $(z', y')$ is chosen for large $\ell$. Since

$$\mu(\text{Per}_j(f)) = 0, \quad \mu(\partial U_N) = 0 \quad \text{and} \quad \mu(U_N) > a_0,$$

if $\varepsilon > 0$ has been chosen small enough, we can neglect pairs $(f^{\ell_i}(x_0), f^{\ell_i}(x_0))$ with $f^{\ell_i}(x_0) \notin W_j$ and small pairs $(f^{\ell_i}(x_0), f^{\ell_i}(x_0))$ such that either $f^{\ell_i}(x_0)$ or $f^{\ell_i}(x_0)$ does not belong to $U_N$ for all $\ell$ sufficiently large. Therefore we may assume that $z' \in W_j$ and if $(z, y)$ is small enough then $(z, y) \subseteq U_N$. When $1 \leq i < i' \leq \zeta_0(\ell)$, $1 \leq i \leq N$, we have

$$\ell_i \leq t_i < u_i \leq k_i \leq \ell_i' \leq t_i' < u_i' \leq k_i'. \tag{4.11}$$

Set $\bar{t}_j = \ell_j$, $\bar{k}_j = k_j$, $\bar{t}_j = t_j$, and $\bar{u}_j = u_j$, to simplify the notation, and suppose that there are pairs $(f^{\ell_i}(x_0), f^{\ell_i}(x_0))$, $1 \leq j < \zeta$, such that

$$f^{\ell_i}(x_0) = \cdots = f^{\ell_i}(x_0) \quad \text{and} \quad \bar{u}_1 \leq \cdots \leq \bar{u}_\zeta.$$

for some $2 \leq \zeta \leq \zeta_0(\ell)$. Then let us claim that $\zeta \leq N$. In fact, if $\zeta > N$, by the choice of $(z', y')$, we can find $1 \leq t_0 \leq N$ and at least two pairs $(z_1, y_1)$ and $(z_2, y_2)$ among $\{(f^j(x_0), f^{k_j}(x_0)) : 1 \leq j \leq \zeta\}$ both contained in $\tilde{A}_\alpha$ such that two strings $(z_1, y_1; f)$ and $(z_2, y_2; f)$ contain $f^{i_1}(x_0) \notin \{z_1, y_1, z_2, y_2\}$, having nonempty intersection except their end points and contradicting our choice with (4.11). Take pairs in $\{(f^{i_k}(x_0), f^{k_j}(x_0)) : 1 \leq i \leq \zeta_0(\ell)\}$ from which $(f^{i_1}(x_0), f^{j_1}(x_0))$ with $f^{i_1}(x_0) \in W_j$ are chosen and reindex if necessary keeping $\ell_i < \ell_{i'}$ when $i < i'$, to define

$$\{(f^{i_k}(x_0), f^{k_j}(x_0)) : 1 \leq i \leq \zeta(\ell)\}$$

with $\zeta(\ell) \leq \zeta_0(\ell)$ as a set of such sufficiently small $(U, \varepsilon)$-quasi-closable pairs $(z, y)$ as above contained in some $\tilde{A}$, with $i \notin I_N$ appeared in $\{f^j(x_0) : 1 \leq j \leq \ell\} \cap A_N \cap U_N$. Then by the observation above we can write

$$\{(z_j', y_j') : 1 \leq j \leq \zeta'(\ell)\} = \{(f^{i_k}(x_0), f^{j_1}(x_0)) : 1 \leq i \leq \zeta(\ell)\}$$

(4.12) with $\zeta'(\ell) \geq \zeta(\ell)/N$ for all $\ell$ sufficiently large, satisfying $z_j' \neq z_{j'}'$ if $j \neq j'$. Thus we have proved that the number of pairs $(z', y') \in \{(f^{i_1}(x_0), f^{j_1}(x_0)) : 1 \leq i \leq \zeta(\ell)\}$ satisfying the following property is $\geq \zeta(\ell)/N$:

- If $(z', y')$ is a sufficiently small $(U, \varepsilon)$-quasi-closable pair, then $z' \in \Sigma_j(U, \varepsilon)$.

Now, consider (4.2) as $\mu(\Sigma_j(U, \varepsilon')) = 1$ for $\varepsilon' > 0$ much smaller than $\varepsilon_0$. Then observe that almost all $(U, \varepsilon)$-quasi-closable pairs $(z, y)$ in $\{f^j(x_0) : 1 \leq j \leq \ell\} \cap A_N \cap U_N$ are actually sufficiently small ones for large $\ell$. In fact, they are qualified to be sufficiently small $(U, \varepsilon)$-quasi-closable pairs if their sizes $\|z - y\|_1$ are small enough (see the previous Remark below Lemma 4.4), which can be arbitrarily small according to the choice of small $\varepsilon' > 0$. On the other hand, from (4.8), (4.10) and Lemma 4.6, it follows that

$$\mu(A_N \cap U_N) > \frac{a_0}{2} \left(1 - \frac{1}{2^m}\right).$$

(4.13)

We may assume that $\mu(A_i \cap U_N) > 0$ for all $i \notin I_N$. For every $i \notin I_N$ let $\{j_i : 1 \leq i \leq k(i, \ell)\}$ be the set of positive integers in $\{1, \ldots, \ell\}$ such that $f^{j_i}(x_0) \in A_i \cap U_N$ for $i \notin I_N$ indexed as $j_i < j_{i'}$ if $i < i'$. Then, observe that there exists $L(\ell) \geq 1$ such that at least the half of $i \in \{1, \ldots, k(i, \ell)\}$ satisfy

$$j_{i+1} - j_i \leq L(\ell)$$

(4.14) for all $\ell$ sufficiently large. In fact, taking $L(\ell) \geq 3/\mu(A_i \cap U_N)$ and assuming that (4.14) is not true, then we can choose $x_0 \in S(f)$ such that

$$1 = \mu(S(f)) = \lim_{\ell \to +\infty} \frac{1}{\ell} \#\{1 \leq j \leq \ell : f^{j_1}(x_0) \in S(f)\}$$

$$\geq \lim_{\ell \to +\infty} \frac{1}{2\ell} \left(\#\{1 \leq j \leq \ell : f^{j_1}(x_0) \in A_i \cap U_N\} - 1\right) L(\ell)$$

$$\geq \frac{1}{2} \mu(A_i \cap U_N) \frac{3}{\mu(A_i \cap U_N)} > 1,$$

exhibiting a contradiction. Hence, given $0 < \varepsilon_0 \leq \varepsilon$ if $\varepsilon' > 0$ is small enough depending on

$$L = \max_{1 \leq i \leq N} L(\ell),$$

we always have property (E-c) with $\varepsilon$ replaced by $\varepsilon_0$ in the selection process of Lemma 4.4 through Lemma 4.1. Then, by Lemma 4.4, we can finally obtain sufficiently small $(U, \varepsilon)$-quasi-closable pairs satisfying also property (E-b) in the string
Then, it is easy to see that there exist a subsequence $n_i$ for every $n_i \in I$ such that at least $4mN$ times the frequency of points of $\{f^j(x_0) : 0 \leq j \leq \ell\}$ in $U_N$ hitting $\Sigma_j(U, \varepsilon_0)$ within positive $L$-iterates is more than or equal to that of those hitting $\Sigma_j(U, \varepsilon)$ for large $\ell$, which amounts to (4.13). Let $\varepsilon_n > 0$, $n = 1, 2, \ldots$, be a sequence of positive integers converging to 0. Then the above investigation with $\varepsilon_0 = \varepsilon_n$ implies

$$
\mu\left(\bigcup_{j=0}^{L} f^{-j}(\Sigma_j(U, \varepsilon_n)) \cap U_N\right)
= \lim_{\ell \to +\infty} \frac{1}{\ell} \#\{1 \leq j \leq \ell : f^j(x_0) \in \bigcup_{j=0}^{L} f^{-j}(\Sigma_j(U, \varepsilon_n)) \cap U_N\}
\geq \lim_{\ell \to +\infty} \frac{1}{4mN\ell} \#\{1 \leq j \leq \ell : f^j(x_0) \in \Sigma_j(U, \varepsilon) \cap \Sigma_j(U, \varepsilon) \cap U_N\}
= \frac{1}{4mN} \mu(\Sigma_j(U, \varepsilon)) \cap U_N \geq \frac{a_0}{8mN} \left(1 - \frac{1}{2m}\right)
$$

for every $n \geq 1$. For $n \geq 1$ set

$$
B_n = \bigcup_{l \geq n} \bigcup_{j=0}^{L} f^{-j}(\Sigma_j(U, \varepsilon_l)).
$$

Then, it is easy to see that there exist a subsequence $n_i \to +\infty$, $i = 1, 2, \ldots$, and $i_0 \geq 1$ such that

$$
\mu\left(\bigcap_{l \geq i_0} B_{n_i} \cap U_N\right) > \frac{a_0}{16mN} \left(1 - \frac{1}{2m}\right).
$$

Since

$$
\bigcap_{i \geq i_0} B_{n_i} = \bigcap_{i \geq i_0} \bigcup_{l \geq n_i, j=0}^{L} f^{-j}(\Sigma_j(U, \varepsilon_l)) = \bigcup_{j=0}^{L} \bigcap_{i \geq i_0} \bigcup_{l \geq n_i} \Sigma_j(U, \varepsilon_l) \subset \Sigma^{nu}_j(U, f),
$$

we have

$$
\mu(\Sigma^{nu}_j(U, f) \cap U_N) > \frac{a_0}{16mN} \left(1 - \frac{1}{2m}\right).
$$

This contradicts (4.9) and proves (4.1) through (4.7), completing the proof of Theorem A.

REFERENCES

[1] C. Bonatti, L. J. Díaz and M. Viana, Dynamics Beyond Uniform Hyperbolicity, A Global Geometric and Probabilistic Perspective, Encyclopaedia of Mathematical Sciences, 102. Mathematical Physics, III. Springer-Verlag, Berlin, 2005.

[2] R. Bowen, Entropy expansive maps, Trans. Amer. Math. Soc., 164 (1972), 323–331.

[3] Y. T. Cao, D. W. Yang and Y. L. Zang, The entropy conjecture for dominated splitting with multi 1D centers via upper semi-continuity of the metric entropy, Nonlinearity, 30 (2017), 3076–3087.

[4] Y. L. Cao and D. W. Yang, On Pesin’s entropy formula for dominated splittings without mixed behavior, J. Diff. Equ., 261 (2016), 3964–3986.

[5] A. Castro, The ergodic closing lemma for nonsingular endomorphisms, preprint, (2009), arXiv:0906.2031v2.

[6] S. Crovisier, Partial hyperbolicity far from homoclinic bifurcations, Advances in Math., 226 (2011), 673–726.
A FORWARD ERGODIC CLOSING LEMMA AND THE ENTROPY CONJECTURE

[7] J. Franks, Necessary conditions for stability of diffeomorphisms, Trans. Amer. Math. Soc., 158 (1971), 301–308.

[8] J. Franks and M. Misiurewicz, Topological methods in dynamics, Handbook of dynamical systems, North-Holland, Amsterdam, 1A (2002), 547–598.

[9] S. Hayashi, An extension of the ergodic closing lemma, Ergod. Th. Dynam. Sys., 30 (2010), 773–808.

[10] M. W. Hirsch, J. Palis, C. C. Pugh and M. Shub, Neighborhoods of hyperbolic sets, Inventiones Math., 9 (1969/70), 121–134.

[11] M. W. Hirsch, C. C. Pugh and M. Shub, Invariant Manifolds, Lecture Notes in Mathematics, Vol. 583. Springer-Verlag, Berlin-New York, 1977.

[12] G. Liao, M. Viana and J. G. Yang, The entropy conjecture for diffeomorphisms away from tangencies, J. Eur. Math. Soc., 15 (2013), 2034–2060.

[13] P. D. Liu and K. N. Lu, A note on partially hyperbolic attractors: Entropy conjecture and SRB measures, Discrete Contin. Dyn. Syst., 35 (2015), 341–352.

[14] R. Mañé, An ergodic closing lemma, Ann. of Math., 116 (1982), 503–540.

[15] R. Mañé, Ergodic Theory and Differentiable Dynamics, Ergebnisse der Mathematik und Ihrer Grenzgebiete, 8. Springer-Verlag, Berlin, 1987.

[16] M. Misiurewicz, Diffeomorphisms without any measure of maximal entropy, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys., 21 (1973), 903–910.

[17] M. Misiurewicz, Topological conditional entropy, Studia Math., 55 (1976), 175–200.

[18] K. Moriyasu, The ergodic closing lemma for $C^1$ regular maps, Tokyo J. Math., 15 (1992), 172–183.

[19] V. Pliss, A hypothesis due to Smale, Diff. Eq., 8 (1972), 203–214.

[20] C. C. Pugh, The closing lemma, Amer. J. Math., 89 (1967), 956–1009.

[21] C. C. Pugh and C. Robinson, The $C^1$ closing lemma, including Hamiltonians, Ergod Th. Dynam. Sys., 3 (1983), 261–313.

[22] M. Qian and Z. S. Zhang, Ergodic theory for axiom A endomorphisms, Ergod. Th. Dynam. Sys., 15 (1995), 161–174.

[23] D. Ruelle and D. Sullivan, Currents, flows and diffeomorphisms, Topology, 14 (1975), 319–327.

[24] R. Saghin and Z. H. Xia, The entropy conjecture for partially hyperbolic diffeomorphisms with 1-D center, Topology Appl., 157 (2010), 29–34.

[25] M. Shub, Dynamical systems, filtrations and entropy, Bull. Amer. Math. Soc., 80 (1974), 27–41.

[26] M. Shub and R. F. Williams, Entropy and stability, Topology, 14 (1975), 329–338.

[27] M. Urbanski and C. Wolf, SRB measures for Axiom A endomorphisms, Math. Res. Lett., 11 (2004), 785–797.

[28] P. Walters, An Introduction to Ergodic Theory, Graduate Texts in Mathematics, 79. Springer-Verlag, New York-Berlin, 1982.

[29] L. Wen, The $C^1$ closing lemma for non-singular endomorphisms, Ergod. Th. Dynam. Sys., 11 (1991), 393–412.

[30] L. Wen, Generic diffeomorphisms away from homoclinic tangencies and heterodimensional cycles, Bull. Braz. Math. Soc. (N.S), 35 (2004), 419–452.

[31] Y. Yomdin, Volume growth and entropy, Israel J. Math., 57 (1987), 285–300.

Received June 2019; revised November 2019.

E-mail address: shuhei@ms.u-tokyo.ac.jp