ERGODIC OPTIMIZATION THEORY FOR A CLASS OF TYPICAL MAPS

WEN HUANG, ZENG LIAN, XIAO MA, LEIYE XU, AND YIWEI ZHANG

Abstract. In this article, we consider the weighted ergodic optimization problem of a class of dynamical systems $T : X \to X$ where $X$ is a compact metric space and $T$ is Lipschitz continuous. We show that once $T : X \to X$ satisfies both the Anosov shadowing property (ASP) and the Mañé-Conze-Guivarc’h-Bousch property (MCGBP), the minimizing measures of generic Hölder observations are uniquely supported on a periodic orbit. Moreover, if $T : X \to X$ is a subsystem of a dynamical system $f : M \to M$ (i.e. $X \subset M$ and $f|_X = T$) where $M$ is a compact smooth manifold, the above conclusion holds for $C^1$ observations.

Note that a broad class of classical dynamical systems satisfy both ASP and MCGBP, which includes Axiom A attractors, Anosov diffeomorphisms and uniformly expanding maps. Therefore, the open problem proposed by Yuan and Hunt in [YH] since 1999 is completely solved consequentially.

1. Introduction

Ergodic optimization theory gives expression to the principle of least action in dynamical systems, and has strong connection with other fields, such as Aubry-Mather theory [Co2, Ma, CIPP] in Lagrangian Mechanics; ground state theory [BLL] in thermodynamics formalism and multifractal analysis; and controlling chaos [HO1, OGY, SGOY] in control theory. In this paper, we will study the typical optimization problem in weighted ergodic optimization theory.

Let $(X, T)$ be a dynamical system, that is, $(X, d)$ is a compact metric space and $T : X \to X$ is a continuous map. Denote by $\mathcal{M}(X, T)$ the set of all $T$-invariant Borel probability measures on $X$, which is a non-empty convex and compact topological space with respect to weak* topology. Denote by $\mathcal{M}^e(X, T) \subset \mathcal{M}(X, T)$ the ergodic measures, which is the set of the extremal points of $\mathcal{M}(X, T)$.

---

Huang is partially supported by NSF of China (11431012, 11731003). Lian is partially supported by NSF of China (11725105, 11671279). Xu is partially supported by NSF of China (11801538, 11871188). Zhang is partially supported by NSF of China (11701200, 11871262).
Let $u : X \to \mathbb{R}$ and $\psi : X \to \mathbb{R}^+$ be continuous functions. The quantity $\beta(u; \psi, X, T)$ defined by

\[
\beta(u; \psi, X, T) := \min_{\nu \in \mathcal{M}(X, T)} \frac{\int u \, d\nu}{\int \psi \, d\nu},
\]

is called the *ratio minimum ergodic average*, and any $\nu \in \mathcal{M}(X, T)$ satisfying

\[
\int u \, d\nu = \int \psi \, d\nu = \beta(u; \psi, X, T)
\]

is called a $(u, \psi)$-*minimizing measure*. Denote that

\[
\mathcal{M}_{\text{min}}(u; \psi, X, T) := \left\{ \nu \in \mathcal{M}(X, T) : \int u \, d\nu = \beta(u; \psi, X, T) \right\}.
\]

By compactness of $\mathcal{M}(X, T)$, and the continuity of the operator $\int u \, d\cdot$, it directly follows that $\mathcal{M}_{\text{min}}(u; \psi, X, T) \neq \emptyset$, which contains at least one ergodic $(u, \psi)$-minimizing measure by ergodic decomposition. The purpose of this paper is to establish generic properties of observation functions on the minimizing measures of a considerably broad class of dynamical systems which covers most typical cases such as *Axiom A attractors*, *Anosov diffeomorphisms* and *non-invertible uniformly expanding maps*. Precisely, systems considered in this paper are assumed to be Lipschitz and satisfy the so-called *Anosov Shadow property* (abbr. ASP) and *Mañé-Conze-Guivarc’h-Bousch property* (abbr. MCGBP). In the existing literature, *Axiom A attractors*, *Anosov diffeomorphisms* and *non-invertible uniformly expanding maps* all satisfy both ASP and MCGBP. For the sake of completeness, we give a brief proof of Anosov diffeomorphisms satisfying MCGBP in the Appendix Section A, while ASP is a standard property of Anosov diffeomorphisms thus the proof of which is not repeated in this paper.

For each real number $\eta \geq 0$, we call a sequence $\{x_i\}_{i=0}^{n-1} \subset X$ a periodic $\eta$-*pseudo-orbit* of $(X, T)$, if each $x_{i+1}$ belongs to an $\eta$-neighbourhood of $T(x_i)$, for all $i = 0, \ldots, n - 1 \mod n$. With this convention, we say that $(X, T)$ satisfies Anosov Shadow property (abbr. ASP) and Mañé-Conze-Guivarc’h-Bousch property (abbr. MCGBP) if

**A. (ASP)** There are positive constants $\lambda, \delta, C, L$ such that

1. For $n \in \mathbb{N}$ and $x, y \in X$ with $d(T^i x, T^i y) \leq \delta$ for all $0 \leq i \leq n$, one has for all $0 \leq k \leq n$, $d(T^k x, T^k y) \leq Ce^{-\lambda \min(k, n-k)}(d(x, y) + d(T^n x, T^n y))$.

2. For any $0 \leq \eta \leq \delta$, $n \geq 1$ and a periodic $\eta$-pseudo-orbit $\{x_i\}_{i=0}^{n-1}$, there is a periodic orbit $\{T^i x\}_{i=0}^{m-1}$ with period $m$ such that $m | n$ and $d(x_i, T^i x) \leq L\eta$, $\forall 0 \leq i \leq n - 1$.

**B. (MCGBP)** For any $0 < \alpha \leq 1$, there exists positive integer $K = K(\alpha)$ such that for all $u \in C^{0, \alpha}(X)$, there is $v \in C^{0, \alpha}(X)$ such that

\[
u := u_K - v \circ T^K + v - \beta(u; X, T) \geq 0
\]
where \( u_K = \frac{1}{K} \sum_{i=0}^{K-1} u \circ T^i \) and \( \beta(u; X, T) = \min_{\nu \in \mathcal{M}(X, T)} \int ud\nu \).

Here, \( \alpha \in (0, 1] \) and \( C^{0,\alpha}(X) \) is the space of \( \alpha \)-Hölder continuous real-valued function on \( X \) endowed with the \( \alpha \)-Hölder norm \( \| u \|_\alpha := \| u \|_0 + [u]_\alpha \), where \( \| u \|_0 := \sup_{x \in X} |u(x)| \) is the super norm, and \( [u]_\alpha := \sup_{x \neq y} \frac{|u(x) - u(y)|}{\alpha d(x, y)} \). Also note that when \( \alpha = 1 \), \( C^{0,1}(X) \) becomes the collection of all real valued Lipschitz continuous functions, and \( [u]_1 \) becomes the minimum Lipschitz constant of \( u \).

In summary, let \( \mathcal{C} \) be the set of triple \((X, T, \psi)\) satisfying the following properties:

- H1) \((X, d)\) is a compact metric space and \( T : X \to X \) is Lipschitz continuous;
- H2) \((X, T)\) satisfying ASP and MCGBP;
- H3) \( \psi : X \to \mathbb{R}^+ \) are continuous.

The main results obtained in this paper is summarized in the following:

**Theorem 1.1.** Suppose \((X, T, \psi)\) \(\in\) \(\mathcal{C}\), then the following hold:

I) For \( \alpha \in (0, 1] \), if \( \psi \in C^{0,\alpha}(X) \), then there exists an open and dense set \( \mathcal{P} \subset C^{0,\alpha}(X) \) such that for any \( u \in \mathcal{P} \), \((u, \psi)\)-minimizing measure is uniquely supported on a periodic orbit of \( T \).

II) If \((X, T)\) is a sub-system of a dynamical system \((M, f)\) (i.e. \( X \subset M \) and \( T = f|_X \)) and \( \psi \in C^{0,1}(M) \), where \( M \) is a compact \( C^\infty \) Riemannian manifold, then there exists an open and dense set \( \mathcal{P} \subset C^{1,0}(M) \) such that for any \( u \in \mathcal{P} \), the \((u|_X, \psi|_X)\)-minimizing measure of \((X, T)\) is uniquely supported on a periodic orbit of \( T \), where \( C^{1,0}(M) \) is the Banach space of continuous differentiable functions on \( M \) endowed with the standard \( C^1 \)-norm.

**Remark 1.2.** It is worth to point out that Theorem 1.1 only requires \((X, T)\) satisfy ASP and MCGBP, which means that, in particular, neither topological transitivity for Anosov diffeomorphisms (although it is conjectured that Anosov diffeomorphisms are always topological transitive) nor non-wandering property for Axiom A attractors are needed. If in addition \( X \supset \text{supp}(\mu) \) for all \( \mu \in \mathcal{M}(M, f) \) in Theorem 1.1 (II), then for any \( u \in \mathcal{P} \), the \((u, \psi)\)-minimizing measure of \((M, f)\) is also uniquely supported on a periodic orbit of \( f \).

When \( \psi \equiv 1 \), our theoretic model is reduced into the framework of the classical ergodic optimization theory. In this case, as Theorem 1.1 being applicable on both Axiom A attractors and non-invertible expanding maps, part II) of Theorem 1.1 completely solves the open problem which is originally proposed by Yuan and Hunt in [YH, Conjecture 1.1]. Yuan and Hunt’s conjecture provides a mathematical mechanism on Hunt and Ott’s experimental and heuristic results in [HO2, HO3], and becomes one of the fundamental questions raised in the field of ergodic optimization theory,
which has attracted sustained attentions and yielded considerable results, for instances [BZ, Bo1, Bo2, Bo4, Co1, CLT, Mo, QS]. For a more comprehensive survey for the classical ergodic optimization theory, we refer the reader to Jenkinson [Je1, Je2], to Bochi [B], to Baraviera, Leplaideur, Lopes [BLL], and to Garibaldi [Ga] for a historical perspective of the development in this area. In the existing literature, the best result towards to Yuan and Hunt’s conjecture is obtained by Contreras [Co1]. The main result obtained in [Co1] states that for a uniformly expanding map the minimizing measures of (topological) generic Lipschitz observations are uniquely supported on periodic orbits, which is clearly a straightforward consequence of Part I) of Theorem 1.1.

On the other hand, the reason of adding the nonconstant weight $\psi$ mainly lies in the studies on the zero temperature limit (or ground state) of the $(u, \psi)$-weighted equilibrium state for thermodynamics formalism, i.e., the measure $\mu_{u, \psi} \in \mathcal{M}(X, T)$, which satisfies

$$\mu_{u, \psi} := \arg \max \left\{ h_\nu(T) + \int \frac{ud\nu}{\psi d\nu} : \forall \nu \in \mathcal{M}(X, T) \right\},$$  \hspace{1cm} (1.2)

where $h_\nu$ is the Kolmogorov-Sinai entropy of $\nu$. Such weighted equilibrium state arises naturally in the studies of non-conformal multifractal analysis (e.g. high dimensional Lyapunov spectrum) for asymptotically (sub)additive potentials, see works [BF, BCW, FH]. In fact, when the ground state exists, (i.e., the limit $\lim_{t \to +\infty} \mu_{tu, \psi}$ exists), then the limit formulates a special candidate of $(-u, \psi)$-minimizing measure.

The structure of the paper is organized as follows. In Section 2, we give the proof of Theorem 1.1. In Section 3, we consider the case of observation functions with high regularity, for which some partial results and remaining questions are presented; In Appendix A, we briefly explain why Anosov diffeomorphisms being MCGBP for the sake of completeness.

2. Proof of Theorem 1.1

Before starting the proof, we introduce some notions at first for the sake of convenience. For $(X, T, \psi) \in \mathcal{C}$ and a continuous function $u : X \to \mathbb{R}$, define

$$Z_{u, \psi} := \bigcup_{\mu \in \mathcal{M}_{\min}(u; \psi, X, T)} \text{supp}(\mu).$$  \hspace{1cm} (2.1)

For $\alpha \in (0, 1]$, a non-empty subset $Z$ of $X$ and a periodic orbit $O$ of $(X, T)$, define the $\alpha$-deviation of $O$ with respect to $Z$ by

$$d_\alpha, Z(O) = \sum_{x \in O} d^\alpha(x, Z),$$
where we recall that $d$ is the metric on $X$. Let $\lambda, \delta, C, L$ be the constants as in ASP and fix these notations. Define $D : X \times X \to [0, +\infty)$ by

$$D(x, y) = \begin{cases} \delta, & \text{if } d(x, y) \geq \delta, \\ d(x, y), & \text{if } d(x, y) < \delta. \end{cases}$$

By a periodic orbit $\mathcal{O}$ of $(X, T)$, the gap of $\mathcal{O}$ is defined by

$$D(\mathcal{O}) = \begin{cases} \delta, & \text{if } \#\mathcal{O} = 1, \\ \min_{x, y \in \mathcal{O}, x \neq y} D(x, y), & \text{if } \#\mathcal{O} > 1. \end{cases} \quad (2.2)$$

We will prove Part I) and Part II) of Theorem 1.1 separately.

2.1. Proof of Part I) of Theorem 1.1 The proof of Part I) of Theorem 1.1 mainly contains two steps:

Step 1. We show how to construct periodic orbit $\mathcal{O}$ of $(X, T)$ to make the ratio $D^\alpha(\mathcal{O}) d^{\alpha, Z}(\mathcal{O})$ as large as needed. Such a periodic orbit will be a candidate to support the minimizing measures of observations nearby $u$.

Step 2. We show that for any given $u \in C_0(X)$ there exists an open set $\mathcal{U}$ of observations being arbitrarily close to $u$, whose minimizing measures are all supported on a single periodic orbit $\mathcal{O}$.

The main aim of Step 1 can be summarized into the following Proposition:

**Proposition 2.1.** Let $(X, T)$ satisfy $H1)$ and ASP, $u : X \to \mathbb{R}$ and $\psi : X \to \mathbb{R}^+$ are continuous. Then for any $\alpha \in (0, 1]$, a given $\hat{L} > 0$ and $T$-forward-invariant non-empty subset $Z \subset X$ (i.e. $T(Z) \subset Z$), there exists an periodic orbit $\mathcal{O}$ of $(X, T)$ such that

$$\frac{D^\alpha(\mathcal{O})}{d^{\alpha, Z}(\mathcal{O})} > \hat{L}. \quad (2.3)$$

Similarly, the main aim of Step 2 can be summarized into the following Proposition:

**Proposition 2.2.** Let $(X, T, \psi) \in \mathcal{C}$ (satisfying $H1)$, H2) and H3)) and $\psi, u \in C^{0, \alpha}(X)$ for some $\alpha \in (0, 1]$. Then for any $\varepsilon > 0$, there exist $\hat{L}, \hat{\delta} > 0$ which depend on $\varepsilon, \alpha, u, \psi$ and system constants only such that the following holds: If there is a periodic orbit $\mathcal{O}$ of $(X, T)$ satisfying

$$\frac{D^\alpha(\mathcal{O})}{d^{\alpha, Z_u, \psi}(\mathcal{O})} > \hat{L}, \quad (2.4)$$

then the observation function $u_{\varepsilon, h} := u + \varepsilon d^\alpha(\cdot, \mathcal{O}) + h$ has a unique minimizing measure

$$\mu_{\mathcal{O}} := \frac{1}{\#\mathcal{O}} \sum_{x \in \mathcal{O}} \delta_x.$$
whenever $h \in C^{0,\alpha}(X)$ with $\|h\|_\alpha < 10\epsilon$ and $\|h\|_0 < \frac{\beta^{n}(\mathcal{O})}{\# \mathcal{O}} \cdot \delta$.

It is clear that the collection of $u_{e,h}$ in Proposition 2.2 forms an non-empty open subset of $C^{0,\alpha}(X)$, which is about $\varepsilon$-apart from $u$. Since $\varepsilon$ can be taken arbitrarily small and the existence of periodic orbit satisfying (2.4) are guaranteed by Proposition 2.1, Part I) of Theorem 1.1 follows. Thus it remains to prove Proposition 2.1 and 2.2.

2.1.1. Proof of Proposition 2.1. At first, we introduce a lemma giving a quantified estimate of the denseness of periodic orbits, which can be viewed as a version of Quas and Bressaud’s periodic approximation lemma.

**Lemma 2.3.** Let $(X,T)$ be a dynamical system satisfying H1) and ASP, $Z$ be a nonempty $T$-forward-invariant subset of $X$. Then for all $\alpha \in (0,1]$ and $k > 0$, one has

$$\lim_{n \to \infty} n^k \min_{\mathcal{O} \in \mathcal{O}^n} d_{\alpha,Z}(\mathcal{O}) = 0,$$

where $\mathcal{O}^n$ denote the collection of all periodic orbits of $(X,T)$ with period not larger than $n$.

**Proof.** We follow the arguments in [BQ]. Before going to the proof, we need to state two technical results first. The following lemma is Lemma 5 of [BQ].

**Lemma 2.4 ([BQ]).** Suppose that $(Y,\sigma)$ is a shift of finite type (with forbidden words of length 2) with $M$ symbols and entropy $h$. Then $(Y,\sigma)$ contains a periodic point of period at most $1 + M e^{(1-h)}$.

**Lemma 2.5.** Let $(X,T)$ be a dynamical system satisfying H1) and ASP, and $Z$ be a nonempty subset of $X$. Then for any $0 < \alpha \leq 1$, $0 < \eta \leq \delta$, $n \geq 0$, and periodic $\eta$-pseudo-orbit $\tilde{O}$ of $(X,T)$ with period $n$, there exists a periodic orbit $O$ of $(X,T)$ with period $m$ such that $m | n$ and

$$d_{\alpha,Z}(O) \leq d_{\alpha,Z}(\tilde{O}_n) + n(L\eta)^\alpha,$$

where $\delta, L$ are the constants as in ASP.

**Proof.** For $n \geq 1$, assume $\tilde{O} = \{x_i\}_{i=0}^{n-1}$ is a periodic $\eta$-pseudo-orbit with period $n$. By ASP, there is a periodic orbit $O = \{x, T x, \cdots, T^{m-1} x\}$ such that $m | n$ and $d(x_i, T^ix) \leq L\eta$ for $0 \leq i \leq n - 1$. Therefore, for $0 < \alpha \leq 1$, one has

$$d_{\alpha,Z}(O) = \sum_{i=0}^{m-1} d^\alpha(T^i x, Z) \leq \sum_{i=0}^{n-1} (d(x_i, Z) + L\eta)^\alpha \leq d_{\alpha,Z}(\tilde{O}) + n(L\eta)^\alpha.$$
This ends the proof. □

Now, we are ready to prove Lemma 2.3

Fix $\alpha, k, Z$ as in the lemma and $\lambda, \delta, C, L$ are the constants as in ASP. Let $\mathcal{P} = \{P_1, P_2, \ldots, P_m\}$ be a finite partition of $X$ with diameter smaller than $\delta$. For $x \in X$, \( \hat{x} \in \{1, 2, 3, \ldots, m\}^\mathbb{N} \) is defined by
\[
\hat{x}(n) = j \text{ whenever } T^n x \in P_j \text{ and } n \in \mathbb{N}.
\]

Denote $\hat{Z} = \{ \hat{x} : x \in Z \}$ and $W_n$ is the collection of length $n$ strings that appear in $\hat{Z}$. Then $K_n := e^{-\frac{\lambda}{2}n}W_n$ grows at a subexponential rate, i.e.
\[
\lim_{n \to \infty} \frac{\log K_n}{n} = 0,
\]
where $h = h_{\text{top}}(\bar{Z}, \sigma)$. Denote
\[
Y_n = \{y_0y_1y_2 \ldots \in W_n^\mathbb{N} : y_i \in W_n \text{ and } y_iy_{i+1} \in W_{2n} \text{ for all } i \in \mathbb{N}\}.
\]

Let $(Y_n, \sigma_n)$ be the 1-step shift of finite type $W_n$. Then $(\bar{Z}, \sigma^n)$ can be considered as a subsystem of $(Y_n, \sigma_n)$. Hence
\[
h_{\text{top}}(Y_n, \sigma_n) \geq h_{\text{top}}(\bar{Z}, \sigma^n) = nh.
\]

Thus from Lemma 2.4, the shortest periodic orbit in $Y_n$ is at most $1 + e^{1-hn}W_n = 1 + eK_n$. Denote one of the shortest periodic orbit in $Y_n$ by $z_1z_2 \cdots z_p z_1z_2 \cdots$ for some $p \leq 1 + eK_n$ and $z_i \in W_n, i = 1, 2, \ldots, p_n$.

Now we construct a periodic pseudo-orbit in $Z$. For $i = 1, 2, \ldots, p_n$, there is $x_i \in Z$ such that the leading length $2n$ string of $\hat{x}_i$ is $z_i z_{i+1}$ (Note $z_{p_n+1} = z_1$). Hence, $T^n x_i$ and $\hat{x}_{i+1}$ have the same leading length $n$ string which implies $d(T^{n+j}x_i, T^jx_{i+1}) < \delta$ (Note $x_{p_n+1} = x_1$) for $j = 0, 1, 2, \ldots, n - 1$. By ASP,
\[
d(T^{n+[\frac{n}{2}]}x_i, T^{[\frac{n}{2}]}x_{i+1}) < Ce^{-\lambda \min([\frac{n}{2}], n-1-[\frac{n}{2}])} \cdot (d(T^nx_i, x_{i+1}) + d(T^{2n-1}x_i, T^{n-1}x_{i+1})) \leq 2\delta Ce^{-\lambda([\frac{n}{2}]-1)}.
\]

Therefore, we select the periodic $2\delta Ce^{-\lambda([\frac{n}{2}]-1)}$-pseudo-orbit $\tilde{O}_n$ in $Z$ with periodic $np_n$ by
\[
\{T^{[\frac{n}{2}]}x_1, T^{[\frac{n}{2}]+1}x_1, \ldots, T^{n+[\frac{n}{2}]-1}x_1, T^{[\frac{n}{2}]}x_2, \ldots, T^{n+[\frac{n}{2}]-1}x_2, T^{[\frac{n}{2}]}x_3, \ldots, T^{n+[\frac{n}{2}]-1}x_{p_n}\}.
\]

By ASP, while $n$ is sufficiently large, we have a periodic orbit $O_n$ with period $m_n$, such that $m_n | np_n$
\[
d_{\alpha, Z}(O_n) \leq d_{\alpha, Z}(\tilde{O}_n) + np_n (2\delta CL)^\alpha e^{-\lambda\alpha([\frac{n}{2}]-1)} = np_n (2\delta CL)^\alpha e^{-\lambda\alpha([\frac{n}{2}]-1)}.
\]
where we used lemma 2.5. Since $p_n \leq 1 + e K_n$ and $K_n$ grows at a subexponential rate, we obtain

$$\limsup_{n \to \infty} n^k \min_{\mathcal{O} \in \mathcal{O}_n} d_{\alpha,Z}(\mathcal{O}) \leq \limsup_{n \to \infty} (\max\{i p_i : 1 \leq i \leq n + 1\})^k \cdot np_n \delta^a e^{-\lambda a ([\frac{n}{2}] - 1)} = 0.$$ 

This ends the proof. □

Let $(X,T)$ satisfy H1) and ASP, $u: X \to \mathbb{R}$ and $\psi: X \to \mathbb{R}^+$ are continuous. Given $\alpha \in (0,1]$, $\tilde{L} > 0$ and a $T$-forward-invariant non-empty subset $Z \subset X$. Now, we are ready to construct the required periodic orbit in Proposition 2.1 satisfying (2.3). Before the rigorous proof, we firstly introduce the idea of the construction in a vague way: One can start with a periodic orbit $O_0$ with long enough period $n$ and a good approximation to $Z$ (say $d_{\alpha,Z}(O_0) < n^{-k}$ for some large $k$); Once the gap of $O_0$, $D(O_0)$, is too small to meet the requirement, $O_0$ can be decomposed into two pseudo periodic orbits, one of which has at most half of the original period $n$; Such pseudo orbits will provide a nearby periodic orbit with same period by ASP, say $O_1$; One can show that the ratio $d_{\alpha,Z}(O_1)$ is bounded by a constant depending on system constants and $\tilde{L}$ only rather than dependending on $d_{\alpha,Z}(O_0)$; Note that the operation of decomposing periodic orbits into periodic orbits with period halved can be done at most $\log_2 n$ times; Therefore, by adjusting the largeness of $n, k$, the above process will end at either a periodic orbit meet the requirement of Proposition 2.1 or a fixed point, both of which will clearly accomplish the proof.

Let $C, L, \lambda, \delta$ be as in ASP. Take $k \in \mathbb{N}$ large enough, on which the condition will be proposed later. By Lemma 2.3 there exists a periodic orbit $O_0$ of $(X,T)$ with period $n$ large enough such that

$$d_{\alpha,Z}(O_0) < \tilde{L}_0 n^{-k} \ll \delta,$$

where $\tilde{L}_0 = 1$. If $D^a(O_0) > \tilde{L} d_{\alpha,Z}(O_0)$, the proof is done. Otherwise, one has that

$$D^a(O_0) \leq \tilde{L} d_{\alpha,Z}(O_0) < \tilde{L}\tilde{L}_0 n^{-k},$$

which is required to be smaller than $\delta^a$ by choosing $n, k$ large enough. Therefore, there are $y \in O_0$ and $1 \leq n_1 \leq n - 1$ such that

$$d(y, T^{n_1}y) < (\tilde{L}\tilde{L}_0)^{\frac{1}{n}} n^{-\frac{k}{n}} < \delta.$$ 

We split the periodic orbit $O_0$ into two pieces of orbit by

$$Q^0_0 = \{y, Ty, \ldots, T^{n_1-1}y\};$$
$$Q^1_0 = \{T^{n_1}y, T^{n_1}y, \ldots, T^{n-1}y\}.$$ 

Note that each of the above segment of orbit induces a $\delta$-pseudo periodic orbit, and moreover, period of one such $\delta$-pseudo periodic orbit does not exceed $\frac{n}{2}$. Without losing any generality, we assume that $n_1 \leq \frac{n}{2}$. 
By ASP, there exists a periodic orbit
\[ \mathcal{O}_1 = \{ z_1, Tz_1, \ldots, T^{m_1-1}z_1 \} \]
such that \( T^{m_1}z = z \), \( m_1|n_1 \) and \( d(T^iz, T^iz_1) \leq L(\tilde{L}\tilde{L}_0)^{\frac{1}{2}}n^{-\frac{k}{2}} \) for all \( 0 \leq i \leq n_1 - 1 \). Therefore, by ASP again, for all \( 0 \leq i \leq n_1 - 1 \),
\[ d(T^iz, T^iz_1) \leq Ce^{-\lambda \min(i, n_1 - i)}2L(\tilde{L}\tilde{L}_0)^{\frac{1}{2}}n^{-\frac{k}{2}}, \]
which \( L(\tilde{L}\tilde{L}_0)^{\frac{1}{2}}n^{-\frac{k}{2}} \) is required to be smaller than \( \delta \), that is, \( L^\alpha(\tilde{L}\tilde{L}_0)n^{-k} < \delta^\alpha \) by choosing \( n, k \) large enough.

Hence,
\[ d_{\alpha, Z}(\mathcal{O}_1) \leq d_{\alpha, Z}(\mathcal{O}_0^0) + \sum_{i=0}^{n_1-1} \left( Ce^{-\lambda \min(i, n_1 - i)}2L \right)^\alpha (\tilde{L}\tilde{L}_0)n^{-k} \]
\[ \leq d_{\alpha, Z}(\mathcal{O}_0) + \frac{2(2CL)^\alpha}{1 - e^{-\lambda \alpha}}(\tilde{L}\tilde{L}_0)n^{-k} \]
\[ < \tilde{L}_1n^{-k}, \]
where \( \tilde{L}_1 = (1 + \frac{2(2CL)^\alpha}{1 - e^{-\lambda \alpha}})\tilde{L}_0 = 1 + \frac{2(2CL)^\alpha}{1 - e^{-\lambda \alpha}} \tilde{L} \).

If \( D^\alpha(\mathcal{O}_1) < \tilde{L}d_{\alpha, Z}(\mathcal{O}_1) \), the proof is done. Otherwise, one repeats the above operation to get another periodic orbit \( \mathcal{O}_2 \) with period \( \leq \frac{n_1}{4} \). Note that, in this case, in order to make the above process repeatable one only need
\[ \tilde{L}\tilde{L}_1n^{-k} < \delta^\alpha \text{ and } L^\alpha \tilde{L}\tilde{L}_1n^{-k} < \delta^\alpha, \]
which is doable by choosing \( n, k \) large enough. Suppose the above operation can be executed \( m \) times resulting at a periodic orbit \( \mathcal{O}_m \). Then, by applying the same argument inductively, one has that
\[ d_{\alpha, Z}(\mathcal{O}_m) < \tilde{L}_mn^{-k}, \]
where \( \tilde{L}_m = (1 + \frac{2(2CL)^\alpha}{1 - e^{-\lambda \alpha}})\tilde{L}_m-1 = \ldots = \tilde{L}_1^m \). Since every operation will (at least) halve the period of the resulting periodic orbit, such process has to end before \( \left( \left\lceil \frac{\log n}{\log 2} \right\rceil + 1 \right) \)-th operation. In order to make each operation doable, one only need \( n, k \) satisfying the following condition
\[ \tilde{L}\tilde{L}_1^{\frac{\log n}{\log 2}+1}n^{-k} = \tilde{L}\tilde{L}_1n^{-k+\frac{\log L_1}{\log 2}} < \delta^\alpha \text{ and } L^\alpha \tilde{L}\tilde{L}_1^{\frac{\log n}{\log 2}+1}n^{-k} = L^\alpha \tilde{L}\tilde{L}_1n^{-k+\frac{\log L_1}{\log 2}} < \delta^\alpha. \]

Note that, after the last operation being executed, there are two possible cases: the resulting periodic orbit either meet the requirement of Proposition 2.1 or is a fixed point of \( T \). In the second case, by the definition of \( D(\mathcal{O}) \) (2.2), requirement of Proposition 2.1 is also met once one additionally choose \( \tilde{L}_1^{\frac{\log n}{\log 2}+1}n^{-k} < \delta^\alpha \), that is, \( \tilde{L}_1n^{-k+\frac{\log L_1}{\log 2}+1} < \delta^\alpha. \)
The proof of Proposition 2.1 is completed.

2.1.2. Proof of Proposition 2.2. Before going to the proof of Proposition 2.2, we need to introduce a technical lemma and some notions that play important roles in later proof.

Lemma 2.6. Let \((X,T)\) be a dynamical system satisfying H1) and MCGBP. Then for all \(0 < \alpha \leq 1\), strictly positive \(\psi \in C^{0,\alpha}(X)\) and \(u \in C^{0,\alpha}(X)\), there is \(v \in C^{0,\alpha}(X)\) such that

1. \(u_K - v \circ T^K + v - \beta(u;\psi,X,T)\psi_K \geq 0\);
2. \(Z_{u,\psi} \subset \{x \in X : (u_K - v \circ T^K + v - \beta(u;\psi,X,T)\psi_K)(x) = 0\}\),

where \(K = K(\alpha)\) is the natural number as in MCGBP and \(Z_{u,\psi}\) is given by (2.1).

Proof. (1). By MCGBP, we only need to show that \(\beta(u - \beta(u;\psi,X,T)\psi;X,T) = 0\), that is,

\[
\min_{\mu \in \mathcal{M}(X,T)} \int u - \beta(u;\psi,X,T)\psi d\mu = 0.
\]

It is immediately from the fact

\[
\min_{\mu \in \mathcal{M}(X,T)} \frac{\int ud\mu}{\int \psi d\mu} = \beta(u;\psi,X,T),
\]

where we used the assumption \(\psi\) is strictly positive.

(2). By a probability measure \(\mu \in \mathcal{M}_{min}(u;\psi,X,T)\), we have

\[
\int u_K - v \circ T^K + v - \beta(u;\psi,X,T)\psi_K d\mu = \int u - \beta(u;\psi,X,T)\psi d\mu = 0.
\]

Combining (1) and the fact \(u_K - v \circ T^K + v - \beta(u;\psi,X,T)\psi_K\) is continuous, we have \(\text{supp}(\mu) \subset \{x \in X : (u_K - v \circ T^K + v - \beta(u;\psi,X,T)\psi_K)(x) = 0\}\). Therefore, by the continuity of \(u_K - v \circ T^K + v - \beta(u;\psi,X,T)\psi_K\), one has that \(Z_{u,\psi} = \bigcup_{\mu \in \mathcal{M}_{min}(u;\psi,X,T)} \text{supp}(\mu) \subset \{x \in X : (u_K - v \circ T^K + v - \beta(u;\psi,X,T)\psi_K)(x) = 0\}\). This ends the proof.

Remark 2.7. For convenience, in the following text, if we need to use lemma 2.6, we use \(\bar{u}\) to represent \(u_K - v \circ T^K + v - \beta(u;\psi,X,T)\psi_K\) for short. Then, \(\bar{u} \geq 0\) and \(Z_{u,\psi} \subset \{x \in X : \bar{u}(x) = 0\}\).

Fix \(\varepsilon, \alpha, \psi, u\) as in Proposition 2.2 \(K, \bar{u}\) as in remark (2.7), and \(C, \delta\) as in ASP. By remark 2.7 one has that \(\bar{u} \geq 0\) and \(Z_{u,\psi} \subset \{x \in X : \bar{u}(x) = 0\}\).
In stead of investigating the minimizing measure of $\bar{u} + \varepsilon d^\alpha(\cdot, \mathcal{O}) + h$, we consider a modified observation $G := \bar{u} + \varepsilon d^\alpha(\cdot, \mathcal{O}) + h - a_{\mathcal{O}} \psi_K$ which will provide more conveniences, where

$$a_{\mathcal{O}} := \frac{\sum_{y \in \mathcal{O}} (\bar{u}(y) + \varepsilon d^\alpha(y, \mathcal{O}) + h(y))}{\sum_{y \in \mathcal{O}} \psi_K(y)}.$$

Clearly $\int G d\mu_{\mathcal{O}} = 0$. Note that, by the definition of $u_K := \frac{1}{K} \sum_{i=0}^{K-1} u \circ T^i$, for all $\mu \in \mathcal{M}(X, T)$, one has that

$$\frac{\int u + \varepsilon d^\alpha(\cdot, \mathcal{O}) + h d\mu}{\int \psi d\mu} = \frac{\int u_K + \varepsilon d^\alpha(\cdot, \mathcal{O}) + h d\mu}{\int \psi_K d\mu}$$

$$= \frac{\int \bar{u} + \varepsilon d^\alpha(\cdot, \mathcal{O}) + h d\mu}{\int \psi_K d\mu} + \beta(u; \psi, X, T)$$

$$= \frac{\int G d\mu}{\int \psi d\mu} + a_{\mathcal{O}} + \beta(u; \psi, X, T),$$

where we recall that $\beta(u; \psi, X, T)$ is the minimum ergodic average given by (1.1).

Then, in order to show that $\mu_{\mathcal{O}} \in \mathcal{M}_{\min}(u + \varepsilon d^\alpha(\cdot, \mathcal{O}) + h; \psi, X, T)$, it is enough to show that $\mu_{\mathcal{O}} \in \mathcal{M}_{\min}(G; \psi, X, T)$. Since $\psi$ is strictly positive and $\int G d\mu_{\mathcal{O}} = 0$, it is enough to show that

$$\int G d\mu > 0 \text{ for all } \mu \in \mathcal{M}^e(X, T) \setminus \{\mu_{\mathcal{O}}\}. \quad (2.8)$$

Denote that

$$L_{\mathcal{O}} = \frac{D^\alpha(\mathcal{O})}{d_{\alpha, Z, u, \psi}(\mathcal{O})}.$$

Thus one has an equivalent statement of Proposition 2.2 under the same setting as the following:

**Lemma 2.8.** There exist $\hat{L}, \hat{\delta} > 0$ which depend on $\varepsilon, \alpha, u, \psi$ and system constants only such that if $L_{\mathcal{O}} > \hat{L}$, $\|h\|_\alpha \leq 10\varepsilon$ and $\|h\|_0 < \frac{D^\alpha(\mathcal{O})}{\hat{L}} \cdot \hat{\delta}$, then (2.8) holds.
Proof. Put Area$_1 := \left\{ x \in X : d(x, \mathcal{O}) \leq \left( \frac{|a_{\mathcal{O}}|\|\psi\|_0 + \|h\|_0}{\varepsilon} \right)^{\frac{1}{\alpha}} \right\}$. Note that

\[ |a_{\mathcal{O}}| = \left| \sum_{y \in \mathcal{O}} (\bar{u}(y) + h(y)) \right| \leq \sum_{y \in \mathcal{O}} \left( \|\bar{u}\|_0 a^a(y, Z_u, \psi) + \|h\|_0 \right) \sum_{y \in \mathcal{O}} \psi_{\min} \]

Hence

\[ \frac{\left( \frac{|a_{\mathcal{O}}|\|\psi\|_0 + \|h\|_0}{\varepsilon} \right)^{\frac{1}{\alpha}}}{D(\mathcal{O})^{\frac{1}{2}}} \leq 2 \left( \frac{\|\bar{u}\|_0 \sum_{y \in \mathcal{O}} \psi_{\min} \left\{ \frac{1}{L_{\mathcal{O}}} + \frac{\|h\|_0}{\varepsilon \psi_{\min} D^a(\mathcal{O})} \right\}^{\frac{1}{\alpha}} \right) \]

Particularly, when $L_{\mathcal{O}} > \frac{2(2 \text{Lip}_T)^a\|\bar{u}\|_0}{\varepsilon \psi_{\min} D^a(\mathcal{O})}$ and $\|h\|_0 < \frac{\varepsilon \psi_{\min}}{2(2 \text{Lip}_T)^a} D^a(\mathcal{O})$, one has

\[ \left( \frac{|a_{\mathcal{O}}|\|\psi\|_0 + \|h\|_0}{\varepsilon} \right)^{\frac{1}{\alpha}} < \frac{D(\mathcal{O})}{2 \text{Lip}_T} \]

where

\[ \text{Lip}_T = \left\{ \begin{array}{ll} 1, & \text{if } \#T(X) = 1, \\ \max \left\{ 1, \sup_{x \neq y} \frac{d(Tx, Ty)}{d(x, y)} \right\}, & \text{if } \#T(X) > 1. \end{array} \right. \]  

Thus in the choice of $L$ and $\delta$, we will require $L \geq \frac{2(2 \text{Lip}_T)^a\|\bar{u}\|_0}{\varepsilon \psi_{\min} D^a(\mathcal{O})}$ and $\delta \leq \frac{\varepsilon \psi_{\min}}{2(2 \text{Lip}_T)^a}$. Then when $L_{\mathcal{O}} > \hat{L}$, and $\|h\|_0 < \frac{D^a(\mathcal{O})}{\varepsilon \text{Lip}_T} \cdot \hat{\delta}$, one has

\[ \left( \frac{|a_{\mathcal{O}}|\|\psi\|_0 + \|h\|_0}{\varepsilon} \right)^{\frac{1}{\alpha}} < \frac{D(\mathcal{O})}{2 \text{Lip}_T}. \]  

Firstly, we show that Area$_1$ contains all $x \in X$ with $G(x) \leq 0$.

Given $x \notin \text{Area}_1$ when Area$_1 \neq X$, we are to show that $G(x) > 0$. There exists $y \in \mathcal{O}$ such that

\[ d(x, y) = d(x, \mathcal{O}) > \left( \frac{|a_{\mathcal{O}}|\|\psi\|_0 + \|h\|_0}{\varepsilon} \right)^{\frac{1}{\alpha}}. \]

Note that

\[ \bar{u} + h - a_{\mathcal{O}}\psi_K \geq h - |a_{\mathcal{O}}|\psi_K \geq -|a_{\mathcal{O}}|\|\psi\|_0 - \|h\|_0 \]

since $\bar{u} \geq 0$ and $\|\psi_K\|_0 \leq \|\psi\|_0$. Then

\[ G(x) = \bar{u}(x) + \varepsilon d^a(x, \mathcal{O}) + h(x) - a_{\mathcal{O}}\psi_K \geq \varepsilon d^a(x, \mathcal{O}) - |a_{\mathcal{O}}|\|\psi\|_0 - \|h\|_0 \]  

\[ \geq \varepsilon d^a(x, \mathcal{O}) - |a_{\mathcal{O}}|\|\psi\|_0 - \|h\|_0. \]
> \varepsilon \cdot \left( \frac{|a_{\mathcal{O}}||\psi||_0 + \|h\|_0}{\varepsilon} \right)^{\frac{1}{\alpha}} - |a_{\mathcal{O}}||\psi||_0 - \|h\|_0 \\
= 0.

Secondly, we will show that by choosing \( L_{\mathcal{O}}, \|h\|_\alpha \) and \( \|h\|_0 \) properly, for any \( z \in X \) which is not a generic point of \( \mu_{\mathcal{O}} \), there is an \( m \in \mathbb{N} \cup \{0\} \) such that \( \sum_{i=0}^{m} G(T^i z) > 0 \). The conditions proposed for \( L_{\mathcal{O}}, \|h\|_\alpha \) and \( \|h\|_0 \) will provide the existence of the constants \( \hat{L} \) and \( \hat{\delta} \) being requested by Proposition 2.2.

Suppose that \( z \in X \) is not a generic point of \( \mu_{\mathcal{O}} \). There are two cases. In the case \( z \notin \text{Area}_1 \) just note \( m = 0 \) since \( G(x) > 0 \) by claim 1.

In the case \( z \in \text{Area}_1 \), there is \( y_0 \in \mathcal{O} \) such that

\[
d(z, y_0) = d(z, \mathcal{O}) \leq \left( \frac{|a_{\mathcal{O}}||\psi||_0 + \|h\|_0}{\varepsilon} \right)^{\frac{1}{\alpha}} < \frac{D(\mathcal{O})}{2 \text{Lip}_T}
\]

by (2.10). If \( d(T^k z, T^k y_0) \leq \delta \) for all \( k \in \mathbb{N} \), by ASP, we have

\[
d(T^k z, T^k y_0) \leq C e^{-\lambda k}(d(z, y_0) + d(T^{2k} z, T^{2k} y_0)) \leq 2 C e^{-\lambda k} \delta \to 0 \text{ as } k \to +\infty.
\]

Then \( z \) must be a generic point of \( \mu_{\mathcal{O}} \) which is impossible by our assumption. Hence, there must be some \( m_0 \in \mathbb{N} \) such that

\[
d(T^{m_0} z, T^{m_0} y_0) > \delta > \frac{D(\mathcal{O})}{2}.
\]

Let \( m_1 \in \mathbb{N} \) be the smallest time such that

\[
\frac{D(\mathcal{O})}{2 \text{Lip}_T} \leq d(T^{m_1} z, T^{m_1} y_0) \leq \frac{D(\mathcal{O})}{2} < \delta.
\]

Then, we have

\[
d(T^{m_1} z, \mathcal{O}) = d(T^{m_1} z, T^{m_1} y_0) \geq \frac{D(\mathcal{O})}{2 \text{Lip}_T}.
\]

Hence, by (2.11), we have

\[
G(T^{m_1} z) = \bar{u}(T^{m_1} z) + \varepsilon \tilde{d}^{\alpha}(T^{m_1} z, \mathcal{O}) + h(T^{m_1} z) - a_{\mathcal{O}} \psi_K(T^{m_1} z) \\
\geq \varepsilon \tilde{d}^{\alpha}(T^{m_1} z, \mathcal{O}) - |a_{\mathcal{O}}||\psi||_0 - \|h\|_0 \\
\geq \varepsilon \cdot \left( \frac{D(\mathcal{O})}{2 \text{Lip}_T} \right)^{\alpha} - |a_{\mathcal{O}}||\psi||_0 - \|h\|_0.
\]

Let \( m_2 \in \mathbb{N} \) the largest time with \( 0 \leq m_2 < m_1 \) such that

\[
T^{m_2} z \in \text{Area}_1.
\]

Then for all \( m_2 < n < m_1 \)

\[
G(T^n z) > 0.
\]
On the other hand, since \( m_2 < m_1 \), by \((2.12)\), one has that for all \( 0 \leq n \leq m_2 \)
\[
d(T^nz, T^ny_0) \leq \frac{D(O)}{2} < \delta.
\]
Then by \( \text{ASP} \), one has that for all \( 0 \leq n \leq m_2 \)
\[
d^\alpha(T^nz, T^ny_0) \leq C^\alpha e^{-\lambda \alpha \min(n, m_2 - n)} \cdot (d(z, y_0) + d(T^{m_2}z, T^{m_2}y_0))^\alpha
\leq 2C^\alpha(e^{-\lambda \alpha n} + e^{-\lambda \alpha (m_2 - n)}) \frac{|a_\alpha| \|\psi\|_0 + \|h\|_0}{\varepsilon}
\]
where we used the assumption that \( z, T^{m_2}z \in \text{Area}_1 \). Therefore,
\[
\sum_{n=0}^{m_2} d^\alpha(T^nz, T^ny_0) \leq \frac{4C^\alpha}{1 - e^{-\lambda \alpha}} \cdot \frac{|a_\alpha| \|\psi\|_0 + \|h\|_0}{\varepsilon}.
\]
Thus, one has that
\[
\sum_{n=0}^{m_2} (G(T^nz) - G(T^ny_0))
= \sum_{n=0}^{m_2} \left( \bar{u}(T^nz) + \varepsilon d^\alpha(T^nz, \mathcal{O}) + h(T^nz) - \bar{u}(T^ny_0) - \varepsilon d^\alpha(T^ny_0, \mathcal{O}) - h(T^ny_0) \right)
+ a_\mathcal{O} \sum_{n=0}^{m_2} (\psi_K(T^ny_0) - \psi_K(T^nz))
\geq \sum_{n=0}^{m_2} \left( \bar{u}(T^nz) - \bar{u}(T^ny_0) + h(T^nz) - h(T^ny_0) + a_\mathcal{O} (\psi_K(T^ny_0) - \psi_K(T^nz)) \right)
\geq - (\|\bar{u}\|_{\alpha} + \|h\|_{\alpha} + |a_\mathcal{O}| \|\psi_K\|_{\alpha}) \sum_{n=0}^{m_2} d^\alpha(T^nz, T^ny_0)
\geq - (\|\bar{u}\|_{\alpha} + \|h\|_{\alpha} + |a_\mathcal{O}| \|\psi_K\|_{\alpha}) \cdot \frac{4C^\alpha}{1 - e^{-\lambda \alpha}} \cdot \frac{|a_\mathcal{O}| \|\psi\|_0 + \|h\|_0}{\varepsilon},
\]
where we used the fact \( d^\alpha(\cdot, \mathcal{O}) \geq 0 \) and \( d^\alpha(T^ny_0, \mathcal{O}) = 0 \). Also note that
\[
|a_\mathcal{O}| = \left| \frac{\sum_{y \in \mathcal{O}} (\bar{u}(y) + h(y))}{\sum_{y \in \mathcal{O}} \psi_K(y)} \right| \leq \frac{\|\bar{u}\|_0 + \|h\|_0}{\psi_{\min}}.
\]
Thus, one has that
\[
\sum_{n=0}^{m_2} (G(T^nz) - G(T^ny_0))
\geq - (\|\bar{u}\|_{\alpha} + \|h\|_{\alpha} + \frac{\|\bar{u}\|_0 + \|h\|_0}{\psi_{\min}} \|\psi_K\|_{\alpha}) \cdot \frac{4C^\alpha}{1 - e^{-\lambda \alpha}} \cdot \frac{|a_\mathcal{O}| \|\psi\|_0 + \|h\|_0}{\varepsilon}
\] \tag{2.15}
by \((2.15)\).
Note that \( m_2 + 1 = p\sharp O + r \) for some nonnegative integer \( p \) and \( 0 \leq r \leq \sharp O - 1 \), then by (2.11), one has that
\[
\sum_{n=0}^{m_2} G(T^m y_0) = \sum_{n=p\sharp O}^{m_2 O + r - 1} G(T^m y_0) \geq -\sharp O \cdot (|a_O||\psi||_0 + \|h\|_0) \tag{2.17}
\]
where we used \( \int G d\mu_O = 0 \).

Now we are ready to estimate \( \sum_{n=0}^{m_1} G(f^m z) \) as the following:
\[
\sum_{n=0}^{m_1} G(T^n z)
\geq \sum_{n=0}^{m_2} G(T^n z) + G(T^{m_1} z)
= \left( \sum_{n=0}^{m_2} (G(T^n z) - G(T^n y_0)) \right) + \left( \sum_{n=0}^{m_2} G(T^n y_0) \right) + G(T^{m_1} z)
\geq -\left( \|u\|_\alpha + \|h\|_\alpha + \|u\|_0 + \|h\|_0 \|\psi\|_K \|_\alpha \right) \cdot \frac{4C^\alpha}{1 - e^{-\lambda_\alpha}} \cdot \frac{|a_O||\psi||_0 + \|h\|_0}{\varepsilon} \tag{by (2.15)}
- \sharp O \cdot (|a_O||\psi||_0 + \|h\|_0) \tag{by (2.17)}
+ \varepsilon \cdot \left( \frac{D(O)}{2L_{ipr}} \right)^\alpha - |a_O||\psi||_0 - \|h\|_0 \tag{by (2.13)}
\geq L_1 D^\alpha(O) - L_2 d_{a,Z_{\psi}}(O) - L_3 \sharp O \|h\|_0
= L_1 D^\alpha(O) \left( 1 - \frac{L_2}{L_1} \frac{1}{L_3 \sharp O} \right) \frac{L_1 D^\alpha(O)}{L_3 \sharp O} \|h\|_0,
\]
where we take \( \|h\|_\alpha \leq 10\varepsilon \) and \( \|h\|_0 \leq \delta^\alpha \), and let\n\[
L_1 = \frac{\varepsilon}{(2L_{ipr})^\alpha},
L_2 = \left( \frac{4C^\alpha(\|u\|_\alpha + 10\varepsilon + \|u\|_0 + \|h\|_0 \|\psi\|_K \|_\alpha)}{(1 - e^{-\lambda_\alpha})\psi_{min}^\varepsilon} + \frac{2\|\psi\|_0}{\psi_{min}} \right) \|u\|_\alpha,
L_3 = \left( \frac{4C^\alpha(\|u\|_\alpha + 10\varepsilon + \|u\|_0 + \|h\|_0 \|\psi\|_K \|_\alpha)}{(1 - e^{-\lambda_\alpha})\psi_{min}^\varepsilon} + \frac{2\|\psi\|_0}{\psi_{min}} \right) (1 + \psi_{min}).
\]
Note that \( L_1, L_2, L_3 \) are positive and depending on \( \varepsilon, u, \psi \) and system constants only.
By taking
\[
L_O > 2 \frac{L_2}{L_1}, \|h\|_0 < \frac{1}{2} \frac{L_1 D^\alpha(O)}{L_3 \sharp O}, \text{ and } m = m_1,
\]
one has that \( \sum_{i=0}^{m} G(T^i z) > 0 \) provided that \( z \) is not a generic point of \( \mu_\mathcal{O} \). Therefore, one possible choice for \( \hat{L} \) and \( \hat{\delta} \) is to let
\[
\hat{L} = \max \left\{ \frac{3L_2}{L_1}, \frac{2(2Lip T)^\alpha \|u\|_\alpha}{\epsilon \psi_{\min}} \right\} \quad \text{and} \quad \hat{\delta} = \min \left\{ 1, \frac{L_1}{3L_3}, \frac{\epsilon \psi_{\min}}{2(2Lip T)^\alpha} \right\}.
\]

Finally, we finish the proof by showing that when \( L_\mathcal{O} > \hat{L}, \|h\|_\alpha \leq 10\varepsilon \), and \( \|h\|_0 \leq \frac{D^\alpha(\mathcal{O})}{\varepsilon} \cdot \hat{\delta} \) with \( \hat{L} \) and \( \hat{\delta} \) given above, the following holds
\[
\int G d\mu \geq 0 \quad \text{for all} \quad \mu \in \mathcal{M}^\varepsilon(X, T).
\]
Given an ergodic probability measure \( \mu \in \mathcal{M}^\varepsilon(X, T) \), in the case \( \mu = \mu_\mathcal{O} \), we have \( \int G d\mu_\mathcal{O} = 0 \). In the case \( \mu \neq \mu_\mathcal{O} \), let \( z \) be a generic point of \( \mu \). Note that \( z \) is not a generic point of \( \mu_\mathcal{O} \). Thus there exists \( m_1 \in \mathbb{N} \) such that
\[
\sum_{n=0}^{m_1} G(T^n z) > 0.
\]
Note that \( T^{m_1+1} z \) is also not a generic point of \( \mu_\mathcal{O} \). Thus we have \( m_1 + 1 \leq m_2 \in \mathbb{N} \) such that
\[
\sum_{n=m_1+1}^{m_2} G(T^n z) > 0.
\]
By repeating the above process, we have \( 0 \leq m_1 < m_2 < m_3 < \cdots \) such that
\[
\sum_{n=m_{i+1}}^{m_{i+1}} G(T^n z) > 0 \quad \text{for} \quad i = 0, 1, 2, 3, \cdots,
\]
where \( m_0 = -1 \). Therefore
\[
\int G d\mu = \lim_{i \to +\infty} \frac{1}{m_{i+1} + 1} \sum_{n=0}^{m_i} G(T^n z)
= \lim_{i \to +\infty} \frac{1}{m_{i+1} + 1} \left( \sum_{n=0}^{m_1} G(T^n z) + \sum_{n=m_1+1}^{m_2} G(T^n z) + \cdots + \sum_{n=m_{i-1}}^{m_i} G(T^n z) \right)
\geq 0.
\]
Hence, \( \mu_\mathcal{O} \in \mathcal{M}_{\min}(u + \varepsilon d^\alpha(\cdot, \mathcal{O}) + h; \psi, X, T) \).

In the end, (2.38) holds (\( \varepsilon \) may need to be modified a bit) by noting that for any \( \varepsilon' > \varepsilon_0 \), the function \( G' := G + (\varepsilon' - \varepsilon) d(\cdot, \mathcal{O}) \) satisfies that
\[
\int (G' - G) d\mu = \int (\varepsilon' - \varepsilon) d(\cdot, \mathcal{O}) d\mu > 0 \quad \forall \mu \in \mathcal{M}^\varepsilon(X, T) \setminus \{\mu_\mathcal{O}\}.
\]
This ends the proof. \( \square \)
So far, we have accomplished the proof of Part I) of Theorem 1.1.

2.2. Proof of Part II) of Theorem 1.1. We will prove the following technical proposition which together with Proposition 2.1 imply the Part II) of Theorem 1.1.

Proposition 2.9. Let \((M, f)\) be a dynamical system on a smooth compact manifold \(M\). Assume that \((X, T)\) is a subsystem of \((M, f)\), which satisfies ASP and MCGBP, and \(T : X \to X\) is Lipschitz continuous. Then for \(0 < \varepsilon < 1\), \(u \in C^{1,0}(M)\) and strictly positive \(\psi \in C^{0,1}(M)\), there exist positive numbers \(\hat{L}_1, \hat{\delta}_1 > 0\) depending on \(\varepsilon, \psi, u\) and system constants only, and \(\hat{\delta}_1' > 0\) depending on \(\psi, u\) and system constants only (independent on \(\varepsilon\)) such that the following hold: if a periodic orbit \(O\) of \((X, T)\) meets the following comparison condition

\[
D(O) > \hat{L}_1 d_{1, Z_u, \psi, T}(O),
\]

(2.18)

then there is a \(w \in C^\infty(M)\) with

\[
\|w\|_0 < \hat{\delta}_1 \varepsilon \quad \text{and} \quad \|D_x w\|_0 < 2 \varepsilon
\]

such that the probability measure

\[
\left\{ \mu_O := \frac{1}{\#O} \sum_{x \in O} \delta_x \right\} = \mathcal{M}_{\min} ((u + w + h)|_X; \psi|_X, X, T),
\]

whenever \(h \in C^{1,0}(M)\) satisfies \(\|D_x h\|_0 < 5 \varepsilon\) and \(\|h\|_0 < \frac{D(O)}{\#O} \cdot \hat{\delta}_1\).

Here \(D_x\) is the derivative of a given function and \(Z_{u, \psi, T}\) is same as the \(Z_{u, \psi}\) given by (2.1) with respect to system \((X, T)\). The reason of adding subindex "\(T\)" is to avoid confusion on notions with such invariant set with respect to system \((M, f)\).

Proof. The proof is based on Proposition 2.2 and the following approximation theorem due to Greene and Wu [GW].

Theorem 2.10. Let \(M\) be a smooth compact manifold. Then \(C^\infty(M) \cap C^{0,1}(M)\) is Lip-dense in \(C^{0,1}(M)\).

In this Theorem, \(C^\infty(M) \cap C^{0,1}(M)\) is Lip-dense in \(C^{0,1}(M)\) means that for any \(g_1 \in C^{0,1}(M)\) and \(\varepsilon > 0\) there is a \(g_2 \in C^\infty\) such that \(\|g_1 - g_2\|_0 < \varepsilon\) and \(\|g_2\|_1 < \varepsilon + \|g_1\|_1\). Especially, \(\|D_x g_2\|_0 < \varepsilon + \|g_1\|_1\).

Fix \(\varepsilon, O, \psi, u\) as in the Proposition 2.9. Note that Proposition 2.2 in the case of \(\alpha = 1\) is applicable for the current setting. Thus, by taking \(\hat{L}_1 = \hat{L}\) and \(O\) satisfying (2.4) for
\( \alpha = 1 \), one has that for any \( g \in C^{0,1}(M) \) satisfying that \( \| g \|_1 \leq 10\varepsilon \) and \( \| g \|_0 < \frac{D(O)}{\varepsilon} \delta \)

\[
\left\{ \mu_O := \frac{1}{\mu O} \sum_{x \in O} \delta_x \right\} = M_{\min} \left( (u + \varepsilon d(\cdot, O) + g)|_{X; \psi|_{X, X, T}} \right),
\]

where \( \hat{L}, \delta \) and \( O \) are as in Proposition 2.2. Denote that

\[
U_{C^{0,1}}(\varepsilon d(\cdot, O)) := \left\{ u + \varepsilon d(\cdot, O) + g \mid g \in C^{0,1}(M), \| g \|_1 \leq 10\varepsilon, \| g \|_0 < \frac{D(O)}{\varepsilon} \delta \right\}.
\]

Note that the only obstacle prevent one to derive Proposition 2.9 from Proposition 2.2 directly is that \( d(\cdot, O) \) is only Lipschitz rather than \( C^1 \). A nature idea to overcome this is to find a \( w \in C^{1,0}(M) \) close to \( \varepsilon d(\cdot, O) \) in \( C^{0,1}(M) \) such that an open neighborhood \( U_{C^{1,0}}(w) \) of \( u + w \) in \( C^{1,0}(M) \) is a subset of \( U_{C^{0,1}}(\varepsilon d(\cdot, O)) \), which is doable by applying Theorem 2.10.

Precisely, for any \( \varepsilon_1 > 0 \), by Theorem 2.10, there exists a function \( w \in C^{\infty}(M) \) such that

\[
\| w \|_1 \leq \| \varepsilon d(\cdot, O) \|_1 + \varepsilon_1
\]

and

\[
\| w - \varepsilon d(\cdot, O) \|_0 < \varepsilon_1.
\]

Therefore,

\[
\| D_x w \|_0 \leq \varepsilon + \varepsilon_1 \quad \text{and} \quad \| w \|_0 \leq \| \varepsilon d(\cdot, O) \|_0 + \varepsilon_1. \tag{2.19}
\]

Next, we choose proper \( \varepsilon_1, \hat{\delta}_1 \) and \( \hat{\delta}'_1 \) to meet the requirement of the proposition as follows. For \( h \in C^{1,0}(M) \) we rewrite \( u + w + h \) as \( u + \varepsilon d(\cdot, O) + (w - \varepsilon d(\cdot, O) + h) \) It remains to make \( w - \varepsilon d(\cdot, O) + h \) satisfying the conditions of \( h \) as in Proposition 2.2 by adjusting \( \varepsilon_1 \). Note that

\[
\| w - \varepsilon d(\cdot, O) + h \|_1 \leq \| w \|_1 + \| \varepsilon d(\cdot, O) \|_1 + \| h \|_1 \leq 2\varepsilon + \varepsilon_1 + \| h \|_1, \tag{2.20}
\]

and

\[
\| w - \varepsilon d(\cdot, O) + h \|_0 \leq \| w - \varepsilon d(\cdot, O) \|_0 + \| h \|_0 < \varepsilon_1 + \| h \|_0. \tag{2.21}
\]

Take

\[
\varepsilon_1 = \min \left\{ \varepsilon, \frac{D(O)}{2\varepsilon O} \cdot \hat{\delta} \right\}, \quad \hat{\delta}'_1 = \text{diam}(M) + 1, \quad \hat{\delta}_1 = \frac{1}{2} \hat{\delta},
\]

and let \( \| h \|_1 < 5\varepsilon \) together with \( \| h \|_0 < \frac{D(O)}{3\varepsilon} \cdot \hat{\delta}_1 \). Then one has that

\[
\| D_x w \|_0 \leq 2\varepsilon \quad \text{and} \quad \| w \|_0 \leq \hat{\delta}_1 \varepsilon \quad \text{by (2.19)}
\]

\[
\| w - \varepsilon d(\cdot, O) + h \|_1 \leq 8\varepsilon < 10\varepsilon \quad \text{by (2.20)}
\]
\[ \|w - \varepsilon d(\cdot, \mathcal{O}) + h\|_0 < \frac{D(O)}{2\varepsilon} \cdot \delta \] by (2.21).

Denote that
\[ \mathcal{U}_{\mathcal{C}^{1,0}}(w) := \left\{ u + w + h \mid h \in \mathcal{C}^{1,0}(M), \|h\|_1 < 5\varepsilon, \|h\|_0 < \frac{D(O)}{2\varepsilon} \cdot \delta \right\}. \]

Thus, \( \mathcal{U}_{\mathcal{C}^{1,0}}(w) \subset \mathcal{U}_{\mathcal{C}^{0,1}}(\varepsilon d(\cdot, \mathcal{O})) \), which is simultaneously a non-empty open subset in \( \mathcal{C}^{1,0}(M) \). This complete the proof. \( \square \)

3. Discussions on the case of \( \mathcal{C}^{s,\alpha} \) Observations

In this section, we consider the case when the observation functions has higher regularity. Unlike the case of \( \mathcal{C}^{0,\alpha} \) and \( \mathcal{C}^{1,0} \) observations, only partial results are presented in this paper. To avoid unnecessarily tedious discussions, we will consider the following model which is relatively simple and illustrative.

Let \( (M, f) \) be a dynamical system on a smooth compact manifold \( M \) and \( \psi : M \to \mathbb{R}^+ \) be a strictly positive continuous function. Denote that \( \text{Per}^{s,\alpha}(M, \psi, f) \) is the collection of function \( u \in \mathcal{C}^{s,\alpha}(M) \) such that \( \mathcal{M}_{\text{min}}(u; \psi, M, f) \) contains only one probability measure which is periodic. Now we define \( \text{Per}^{*}{s,\alpha}(M, \psi, f) \) the collection of function \( u \in \mathcal{C}^{s,\alpha}(M) \) such that \( \mathcal{M}_{\text{min}}(u; \psi, M, f) \) contains at least one periodic probability measure. And \( \text{Loc}^{s,\alpha}(M, \psi, f) \) is defined by
\[ \text{Loc}^{s,\alpha}(M, \psi, f) = \{ u \in \text{Per}^{s,\alpha}(M, \psi, f) : \text{there is } \varepsilon > 0 \text{ such that } \mathcal{M}_{\text{min}}(u + h; \psi, M, f) = \mathcal{M}_{\text{min}}(u; \psi, M, f) \text{ for all } \|h\|_{s,\alpha} < \varepsilon \}. \]

In the case \( s \geq 1 \) and \( \alpha > 0 \) or \( s \geq 2 \), we do not have result like Theorem 1.1. But, we have the following weak version.

**Proposition 3.1.** Let \( f : M \to M \) be a Lipschitz continuous selfmap on a smooth compact manifold \( M \) and \( \psi : M \to \mathbb{R}^+ \) be a strictly positive continuous function. If \( u \in \mathcal{C}(M) \) with \( u \geq 0 \) and there is periodic orbit \( \mathcal{O} \) of \( (M, f) \) such that \( u|_{\mathcal{O}} = 0 \), then for all \( \varepsilon > 0 \), \( s \in \mathbb{N} \) and \( 0 \leq \alpha \leq 1 \), there is a function \( w \in \mathcal{C}^\infty(M) \) with \( \|w\|_{s,\alpha} < \varepsilon \) and a constant \( \rho > 0 \) such that the probability measure
\[ \mu_{\mathcal{O}} = \frac{1}{\#\mathcal{O}} \sum_{x \in \mathcal{O}} \delta_x \in \mathcal{M}_{\text{min}}(u + w + h; \psi, M, f), \]
whenever \( h \in \mathcal{C}^{0,1}(M) \) with \( \|h\|_1 < \rho \) and \( \|h\|_0 < \rho \).

By using proposition 3.1, we have the following result immediately.

**Theorem 3.2.** Let \( f : M \to M \) be a Lipschitz continuous selfmap on a smooth compact manifold \( M \). If \( (M, f) \) has ASP and MCGBP, then \( \text{Loc}^{s,\alpha}(M, \psi) \) is an open dense
subset of \( \text{Per}^{s,\alpha}(M,\psi) \) w.r.t. \( \| \cdot \|_{s,\alpha} \) for integer \( s \geq 1 \), real number \( 0 \leq \alpha \leq 1 \) and \( \psi \in C^{0,1}(M) \) is a strictly positive continuous function.

**Proof.** Immediately from Remark 2.7 and Proposition 3.1. \( \square \)

Without result like Proposition 2.2, we can not get the full result about the generality of \( C^{s,\alpha}(M) \), \( s \geq 2 \) or \( s \geq 1 \) and \( \alpha > 0 \). So we raise the following question:

**Question 3.3.** By a expanding map \( f : \mathbb{T} \to \mathbb{T} : x \to 2x \), is there a \( u \in C^{s,\alpha}(\mathbb{T}) \), \( s \geq 1 \) and \( 0 < \alpha \leq 1 \) or \( s \geq 2 \) such that any function near \( u \) w.r.t. \( \| \cdot \|_{s,\alpha} \) has no periodic minimizing measure?

At last, we complete the proof of proposition 3.1.

**Proof of proposition 3.1.** Fix \( \varepsilon, s, \alpha, O, \psi \) as in proposition. \( C \) and \( \delta \) are the constants as in ASP and \( \text{Lip}_f \) is defined as in (2.9). Just take \( w \in C^\infty \) with \( \| w \|_{s,\alpha} < \varepsilon \), \( w|_O = 0 \) and \( w|_{M\setminus O} > 0 \). For \( 0 \leq r \leq D(O) \), we note

\[
\theta(r) = \min \{ w(x) : d(x, O) \geq r, x \in M \}.
\]

It is clear that \( \theta(0) = 0 \), \( \theta(r) > 0 \) for \( r \neq 0 \) and \( \theta \) is non-decreasing. Now we fix the constants

\[
0 < \rho < \frac{D(O)}{2 \text{Lip}_f}
\]

and positive \( \delta \) smaller than

\[
\min \left\{ \frac{\theta(\rho)\psi_{\min}}{\psi_{\min} + \| \psi \|_0}, \frac{\theta(D(O))}{4C\rho} \cdot \frac{1 - e^{-\lambda}}{2} \cdot \frac{\theta \left( \frac{D(O)}{2 \text{Lip}_f} \right)}{2^{\frac{C\rho\| \psi \|_1}{(1-e^{-\lambda})\psi_{\min}}} + \| O \| + \| O \| \| \psi \|_{\min} + \| \psi \|_{\min} + 1} \right\}.
\]

By fixing \( h \in C^{0,1}(M) \) with \( \| h \|_1 < \delta \) and \( \| h \|_0 < \delta \), we are to show that \( \mu_O \in \mathcal{M}_{\min}(w + h; \psi, M, f) \) which implies that \( \mu_O \in \mathcal{M}_{\min}(u + w + h; \psi, M, f) \) since \( u \geq 0 \) and \( u|_O = 0 \) by assumption.

Note that \( G = w + h - a_O\psi \), where \( a_O := \frac{\sum_{y \in O}(w+h)(y)}{\sum_{y \in O}\psi(y)} \). It is straightforward to see that

\[
|a_O| \leq \frac{\| h \|_0}{\psi_{\min}}, \quad (3.1)
\]

where we used \( w|_O = 0 \). Then \( \int_{\psi d\mu} G d\mu = \int_{\psi d\mu} (w+h) d\mu - a_O \). Therefore, to show that \( \mu_O \in \mathcal{M}_{\min}(w + h; \psi, M, f) \), it is enough to show that

\[
\int G d\mu \geq 0 \quad \text{for all } \mu \in \mathcal{M}(M, f),
\]
where we used the assumption $\psi$ is strictly positive and the fact $\int Gd\mu = 0$.

**Claim 1.** Put $\text{Area}_1 = \{ y \in M : d(y, \mathcal{O}) \leq \rho \}$, then $\text{Area}_1$ contains all $x \in M$ with $G(x) \leq 0$.

**Proof of claim 1.** For $x \notin \text{Area}_1$, we have

\[
G(x) = (w + h - a_\mathcal{O}\psi)(x) \geq \theta(\rho) - |a_\mathcal{O}|\|\psi\|_0 - \|h\|_0 \\
\geq \theta(\rho) - \frac{\|\psi\|_0 + \psi_{\min}\|h\|_0}{\psi_{\min}} \\
> \theta(\rho) - \frac{\|\psi\|_0 + \psi_{\min}}{\psi_{\min}}\|h\|_0 \\
\geq 0.
\]

This ends the proof of claim 1. \hfill \Box

**Claim 2.** If $z \in M$ is not a generic point of $\mu_\mathcal{O}$, then there is $m \in \mathbb{N} \cup \{0\}$ such that $\sum_{i=0}^{m} G(f^i z) > 0$.

**Proof of claim 2.** If $z \notin \text{Area}_1$, just note $m = 0$, we have nothing to prove.

Now we assume that $z \in \text{Area}_1$. There is $y_0 \in \mathcal{O}$ such that $d(z, y_0) = d(z, \mathcal{O}) \leq \rho < \frac{D(\mathcal{O})}{2\text{Lip}_f}$.

If $d(f^k z, f^k y_0) \leq \delta$ for all $k \geq 0$, by ASP, we have

\[
d(f^k z, f^k y_0) \leq Ce^{-\lambda k}(d(z, y_0) + d(f^{2k} z, f^{2k} y_0)) \leq 2Ce^{-\lambda k}\delta \to 0 \text{ as } k \to +\infty.
\]

Hence, $z$ is a generic point of $\mu_\mathcal{O}$ which is impossible by our assumption. Therefore, there must be some $m_1 > 0$ such that $d(f^{m_1} z, f^{m_1} y_0) \geq \delta$. There exists $m_2 > 0$ the smallest time such that

\[
\frac{D(\mathcal{O})}{2\text{Lip}_f} \leq d(f^{m_2} z, f^{m_2} y_0) \leq \frac{D(\mathcal{O})}{2}, \quad (3.2)
\]

where we used the assumption $f$ is Lipschitz with Lipschitz constant $\text{Lip}_f$. Then we have $d(f^{m_2} z, \mathcal{O}) = d(f^{m_2} z, f^{m_2} y_0) \geq \frac{D(\mathcal{O})}{2\text{Lip}_f}$ and

\[
G(f^{m_2} z) = (w + h - a_\mathcal{O}\psi)(f^{m_2} z) \geq \theta \left( \frac{D(\mathcal{O})}{2\text{Lip}_f} \right) - \|h\|_0 - |a_\mathcal{O}|\|\psi\|_0. \quad (3.3)
\]

where we used the definition of $\theta(\cdot)$. On the other hand, $\frac{D(\mathcal{O})}{2\text{Lip}_f} > \rho$ by assumption which implies that

\[
f^{m_2} z \notin \text{Area}_1. \quad (3.4)
\]
We take \( m_3 \) the largest time with \( 0 \leq m_3 \leq m_2 \) such that
\[
f^{m_3}z \in \text{Area}_1,
\]
where we use the assumption \( z \in \text{Area}_1 \). By (3.1), it is clear that \( m_3 < m_2 \) since \( m_2 \) is the smallest time meets (3.2). Then by claim 1,
\[
G(f^n z) > 0 \text{ for all } m_3 < n < m_2.
\] (3.5)

Additionally, by the choice of \( m_2 \) and (3.2) one has that
\[
d(f^n z, f^n y_0) \leq D(O) \leq \delta \text{ for all } 0 \leq n \leq m_3.
\]
Therefore, by ASP, we have for all \( 0 \leq n \leq m_3 \),
\[
d(f^n z, f^n y_0) \leq 2C\rho \frac{1}{1 - e^{-\lambda}}.
\]
Since \( w \geq 0 \) and \( w|O = 0 \), one has
\[
\sum_{n=0}^{m_3} (G(f^n z) - G(f^n y_0)) \geq - (\|h\|_1 + |a_O|\|\psi\|_1) \sum_{n=0}^{m_3} d(f^n z, f^n y_0)
\] (3.6)
\[
\geq - (\|h\|_1 + |a_O|\|\psi\|_1) \cdot \frac{2C\rho}{1 - e^{-\lambda}}.
\]

By assuming that \( m_3 = p\sharp O + q \) for some nonnegative integer \( p \) and \( 0 \leq q \leq \sharp O - 1 \), one has
\[
\sum_{n=0}^{m_3} G(f^n y_0) = \sum_{n=0}^{m_3} G(f^n y_0) \geq -2\sharp O \cdot (\|h\|_0 + |a_O|\|\psi\|_0).
\] (3.7)

where we used the facts \( \int Gd\mu_O = 0 \) and \( G \geq -\|h\|_0 + |a_O|\|\psi\|_0 \). Combining (3.1), (3.3), (3.5), (3.6) and (3.7), we have
\[
\sum_{n=0}^{m_2} G(f^n z) \geq \sum_{n=0}^{m_3} G(f^n z) + G(f^{m_2} z)
\]
\[
= \sum_{n=0}^{m_3} (G(f^n z) - G(f^n y_0)) + \sum_{n=0}^{m_3} G(f^n y_0) + G(f^{m_2} z)
\]
\[
\begin{align*}
\geq & - (\|h\|_1 + |a| \|\psi\|_1) \cdot 2C\rho \left( \frac{1}{1 - e^{-\lambda}} \right) - \mathcal{O} \cdot (\|h\|_0 + |a\| \|\psi\|_0) \\
& + \theta \left( \frac{D(\mathcal{O})}{2C} \right) - \|h\|_0 - |a\| \|\psi\|_0 \\
& = \theta \left( \frac{D(\mathcal{O})}{2C} \right) - \|h\|_1 \cdot \frac{2C\rho}{1 - e^{-\lambda}} \\
& - \left( \frac{2C\rho \|\psi\|_1}{(1 - e^{-\lambda})\psi_{\min}} + \mathcal{O} \frac{\|\psi\|_0}{\psi_{\min}} + \frac{\|\psi\|_0}{\psi_{\min}} + 1 \right) \|h\|_0 \\
& > 0,
\end{align*}
\]

where we used the assumption of \( h \). Therefore, \( m = m_2 \) is the time we need. This ends the proof of claim 2. \( \square \)

Now we end the proof. It is enough to show that for all \( \mu \in \mathcal{M}^c(M, f) \)
\[
\int G d\mu \geq 0.
\]

Given \( \mu \in \mathcal{M}^c(f) \), in the case \( \mu = \mu_{\mathcal{O}} \), it is obviously true. In the case \( \mu \neq \mu_{\mathcal{O}} \), just let \( z \) be a generic point of \( \mu \). Note that \( z \) is not a generic point of \( \mu_{\mathcal{O}} \). By claim 2, we have \( m_1 \in \mathbb{N} \) such that
\[
\sum_{n=0}^{m_1} G(f^n z) > 0
\]
Note that \( f^{m_1+1} z \) is also not a generic point of \( \mu_{\mathcal{O}} \). By claim 2, we have \( m_2 \geq m_1 + 1 \) such that
\[
\sum_{n=m_1+1}^{m_2} G(f^n z) > 0.
\]

By repeating the above process, we have \( 0 \leq m_1 < m_2 < m_3 < \cdots \) such that
\[
\sum_{n=m_1+1}^{m_{i+1}} G(f^n z) > 0, \quad i = 0, 1, 2, 3, \cdots,
\]
where \( m_0 \) is noted by \(-1\). Therefore
\[
\int G d\mu = \lim_{i \to +\infty} \frac{1}{m_i + 1} \sum_{n=0}^{m_i} G(f^n z)
\]
\[
= \lim_{i \to +\infty} \frac{1}{m_i + 1} \left( \sum_{n=0}^{m_1} G(f^n z) + \sum_{n=m_1+1}^{m_2} G(f^n z) + \cdots + \sum_{n=m_i-1}^{m_i} G(f^n z) \right)
\]
\[
\geq 0.
\]

That is, we have \( \mu_{\mathcal{O}} \in \mathcal{M}_{\min}(u + h; \psi, M, f) \) by our beginning discussion. This ends the proof. \( \square \)
Appendix A. Mañé-Conze-Guivarc’h-Bousch’s Property

In this section, we mainly present Bousch’s work (see [Bo3] for detail) to show that uniformly hyperbolic diffeomorphism on a smooth compact manifold has MCGBP. The same argument shows that the Axiom A attractor also has MCGBP.

Let \( f : M \to M \) be a diffeomorphism on a smooth compact manifold \( M \). By a function \( u : M \to \mathbb{R} \) and an integer \( K \geq 1 \), note \( u_K = \frac{1}{K} \sum_{i=0}^{K-1} u \circ f^i \). Since \( f \) is assumed Lipschitz, \( u \in C^{0,\alpha}(M) \) for some \( 0 < \alpha \leq 1 \) implies that \( u_K \in C^{0,\alpha}(M) \). Additionally, one has that
\[
\int ud\mu = \int u_K d\mu \text{ for all } \mu \in \mathcal{M}(M,f).
\]
Therefore,
\[
\beta(u; M, f) = \beta(u_K; M, f) \text{ and } \mathcal{M}_{\min}(u; M, f) = \mathcal{M}_{\min}(u_K; M, f).
\]

Theorem A.1. Let \( f : M \to M \) be an Anosov diffeomorphism on a smooth compact manifold \( M \). Then for \( 0 < \alpha \leq 1 \), there is an integer \( K = K(\alpha) \) such that for all \( u \in C^{0,\alpha}(M) \) with \( \beta(u; M, f) \geq 0 \), there is a function \( v \in C^{0,\alpha}(M) \) such that \( u_K \geq v \circ f^K - v \).

Proof. The proof mainly follows Bousch’s work in [Bo3], to which we refer readers for detailed proof. It is worth to point out that the only difference is that, in our setting, one needs a large integer \( K = K(\alpha) \) to grantee that \( (M, f^K) \) meets the condition in [Bo3], and to replace \( u \) by \( u_K \) as \( \beta(u; M, f) = \beta(u_K; M, f) \). \( \square \)

Acknowledgement

At the end, we would like to express our gratitude to Tianyuan Mathematical Center in Southwest China, Sichuan University and Southwest Jiaotong University for their support and hospitality.

References

[BLL] Baraviera, A.; Leplaideur, R.; Lopes, A.O. Ergodic optimization, zero temperature limits and the max-plus algebra. IMPA, Rio de Janeiro, 2013. (Cited on pages [1] and [2])

[B] Bochi, J. Ergodic optimization of Birkhoff averages and Lyapunov exponents. Proceedings of the International Congress of Mathematicians 2018, Rio de Janeiro, vol. 2, pp. 1821–1842. (Cited on page [3])

[BZ] Bochi, J.; Zhang, Y. Ergodic optimization of prevalent super-continuous functions. Int. Math. Res. Not. 19 (2016), 5988–6017. (Cited on page [4])

[Bo1] Bousch, T. Le poisson n’a pas d’arêtes. Ann. Inst. H. Poincaré Probab. Statist. 36 (2000), no. 4, 489–508. (Cited on page [5])
