HAL
open science

Optimal bounds on correlation decay rates for nonuniform hyperbolic systems

Sandro Vaienti, Hong-Kun Zhang

To cite this version:

Sandro Vaienti, Hong-Kun Zhang. Optimal bounds on correlation decay rates for nonuniform hyperbolic systems. 2016. hal-01312743

HAL Id: hal-01312743

https://hal.science/hal-01312743v1

Preprint submitted on 9 May 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

arXiv:1605.01793v1 [math.DS] 5 May 2016

Optimal bounds on correlation decay rates for nonuniform hyperbolic systems

Sandro Vaienti* Hong-Kun Zhang †

Dedicated to the memory of Nikolai Chernov‡

Abstract

We investigate the decay rates of correlations for nonuniformly hyperbolic systems with or without singularities, on piecewise Hölder observables. By constructing a new scheme of coupling methods using the probability renewal theory, we obtain the optimal bounds for decay rates of correlations for a large class of observables. Our results apply to rather general hyperbolic systems, including Bunimovich Stadia, Bunimovich billiards, semidispersing billiards on a rectangle and billiards with cusps, and to a wide class of nonuniformly hyperbolic maps.

AMS classification numbers: 37D50, 37A25

Contents

1 Introduction **2**
1.1 Motivation and relevant works 2
1.2 Plan of exposition 5

2 Assumptions and main results **6**
2.1 Assumptions 6
2.2 Statement of the main results 11

3 Standard families and the induced map **13**
3.1 Construction of an induced system (F, M, μ_M). 13
3.2 Standard families 15

*Aix Marseille Université, CNRS, CPT, UMR 7332, 13288 Marseille, France and Université de Toulon, CNRS, CPT, UMR 7332, 83957 La Garde, France

†Department of Math. & Stat., University of Massachusetts Amherst, Amherst, MA 01003 hongkun@math.umass.edu

‡A first version of this paper was prepared together with N. Chernov three years ago and a few preliminary results were presented in some conferences. The paper was deeply modified after N. Chernov passed away and we now dedicated it to his memory. We are greatly indebted to the ideas and the suggestions he gave us.

4 Coupling Lemma for the induced map **22**
4.1 Construction of a hyperbolic set $\mathcal{R}^*$ 22
4.2 The Coupling Lemma for the induced map 25

5 Markov tower for the original map **30**
5.1 Construction of the Markov tower for the original map 30
5.2 Proper returns to the base of the tower 33
5.3 Multiple returns to $\mathcal{R}^*$ 35
5.4 Proof of Lemma 5.7 39

6 Coupling Lemma for the original system **45**
6.1 Statement of the Coupling Lemma 45
6.2 Outline of the proof of the Coupling Lemma 48
6.3 Proof of the Coupling Lemma 49

7 Proof of Theorem 1. **56**

8 Proof of Theorem 2.3 and Theorem 3. **61**
8.1 Proof of Theorem 2 61
8.2 Prove of Theorem 3. 63

9 Sufficient conditions for (H2). **64**

10 Applications to hyperbolic systems **69**
10.1 Billiards with property **(H2)(a)**. 70
10.2 Billiards with property **(H2)(b)**. 72
10.3 Application to linked-twist maps 74

11 Appendix. List of major notations **76**

1 Introduction

1.1 Motivation and relevant works

The studies of the statistical properties of 2-dimensional hyperbolic systems with singularities are motivated in large part by mathematical billiards with chaotic behavior, introduced by Sinai in [64] and since then studied extensively by many authors [12, 13, 71, 72, 26].

Statistical properties of chaotic dynamical systems are described by the decay of correlations and by various limiting theorems. Let $(\mathcal{M}, \mathcal{F}, \mu)$ be a *dynamical system*, i.e., a measurable transformation $\mathcal{F}\colon \mathcal{M} \to \mathcal{M}$ preserving a probability measure μ on the Borel sigma algebra of $\mathcal{M}$. For any real-valued functions f and g on $\mathcal{M}$ (often called *observables*) the *correlations* are defined by

$$C_n(f, g, \mathcal{F}) = \int_{\mathcal{M}} (f \circ \mathcal{F}^n)\, g\, d\mu - \int_{\mathcal{M}} f\, d\mu \int_{\mathcal{M}} g\, d\mu \tag{1.1}$$

Note that (1.1) is well defined for all $f, g \in L^2_\mu(\mathcal{M})$. It is a standard fact that $(\mathcal{F}, \mu)$ is *mixing* if and only if

$$\lim_{n\to\infty} C_n(f, g, \mathcal{F}) = 0, \qquad \forall f, g \in L^2_\mu(\mathcal{M}) \tag{1.2}$$

The statistical properties of the system $(\mathcal{M}, \mathcal{F}, \mu)$ is characterized by the *rate of decay of correlations*, i.e., by the speed of convergence in (1.2) for "good enough" functions f and g. If $\mathcal{M}$ is a manifold and $\mathcal{F}$ is a smooth (or piecewise smooth) map, then "good enough" usually means bounded and (piecewise) Hölder continuous.

Generally, mixing dynamical systems (even very strongly mixing ones, such as Bernoulli systems) may exhibit quite different statistical properties, depending on the rate of the decay of correlations. If correlations decay exponentially fast (i.e., $|C_n| = \mathcal{O}(e^{-an})$ with $a > 0$), usually the classical Central Limit Theorem (CLT) holds, as well as many other probabilistic limit laws, such as Weak Invariance Principle (convergence to Brownian motion), which play a crucial role in applications to statistical mechanics; we refer the reader to the surveys in [16, 20, 34, 40, 71] and [26, Chapter 7]. Such strongly chaotic dynamical systems behave very much like sequences of i.i.d. (independent identically distributed) random variables in probability theory.

However, some other mixing and Bernoulli systems have slow rates of the decay of correlations, such as $|C_n| = \mathcal{O}(n^{-a})$. Their statistical properties are usually weak, they exhibit *intermittent* behavior [59]: intervals of chaotic motion are followed by long periods of regular oscillations, etc. Such systems can help to understand the transition from regular to chaotic motion, and for that they have long attracted considerable interest in physics community [42, 54, 69]. We note that if[1] $|C_n| \sim n^{-a}$ with $a \le 1$, then even the classical CLT usually fails. In that case the system can be approximated by an unconventional Brownian motion in which the mean squared displacement grows faster than linearly in time. This may help to explain certain unusual physical phenomena, such as superconductivity and superdiffusion. In particular, for $a = 1$ the mean squared displacement acquires an extra logarithmic factor [7, 22, 66].

An challenging question to ask is "What are the main reasons that have showed down the decay rates of correlations for nonuniformly hyperbolic systems"? It has been a mathematically challenging problem to estimate the rates of the decay of correlations for hyperbolic systems with singularities, including chaotic billiards. The main difficulty is caused by singularities and the resulting fragmentation of phase space during the dynamics, which slows down the global expansion of unstable manifolds. Moreover, the differential of the billiard map is unbounded and has unbounded distortion near the singularities, which aggravates the analysis of correlations: one has to subdivide the vicinity of singularities into countably many "shells" in which distortions can be effectively controlled.

Even for strongly chaotic billiards, exponential upper bounds on correlations were proven only in 1998 when Young [71] introduced her tower construction as

[1] We say $A_n \sim B_n$ if there exist $0 < c_1 < c_2$ such that $c_1 B_n \le A_n \le c_2 B_n$ for all $n \ge 1$.

a universal tool for the description of nonuniformly hyperbolic maps; see also [17]. Young also sharpened her estimates on correlations by combining her tower construction with a coupling technique borrowed from probability [72]. The coupling method was further developed by Bressaud and Liverani in [15] and then reformulated in pure dynamical terms (without explicit tower construction) by Dolgopyat [20, 39] using standard pairs; see also [26, Chapter 7]. Dolgopyat's technique was proven to be efficient in handling various types of strongly chaotic systems with singularities [20, 21, 22, 26, 32, 39, 41].

For weakly chaotic billiards, the first rigorous upper bounds on correlations (based on Young's tower) were obtained in the mid-2000s [55, 28]. The results upper bounds were not optimal, as they included an extra logarithmic factor, which was later removed in [31] by a finer analysis of return time statistics. Besides billiards, the same general scheme for bounding correlations has been applied to linked-twist maps [67] and generalized baker's transformations [10], intermittent symplectic maps [52], and solenoids [2].

Surprisingly little progress has been made in obtaining *lower* bounds on correlations for hyperbolic systems, including billiards. Among rare results in this direction are those for Bunimovich stadia [7] and billiards with cusps [5], where a lower bound was a byproduct of a non-classical CLT that forces correlations to be at least of order $\mathcal{O}(n^{-1})$. For one dimensional non-uniformly expanding maps and for Markov maps, lower bounds on correlations have been derived via the renewal methods by Sarig [63], later improved by Gouëzel [44]. The renewal techniques were then extended to more general non-uniformly expanding maps and certain nonuniformly hyperbolic systems; see [46, 48, 56, 53] and references therein. The main scheme of the renewal methods relies on the construction of an induced map, for which the corresponding transfer operator has a spectral gap on a certain functional space. Actually the main reason why it is so difficult to apply operator technique to billiards and related hyperbolic systems with singularities, is the lack of a suitable functional space on which the transfer operator for the induced map would have a spectral gap (and would be aperiodic).

For chaotic billiards and their perturbations, a suitable Banach space of functions was constructed in [36, 37], and the spectral gap for transfer operators was proven to exist. But it is still difficult to apply the renewal operator methods on these systems because in order to take care of the unbounded differential of the billiard map, the norms defined in [36, 37] cannot directly produce the necessary estimates needed for the renewal technique [63]. We should however stress that similar difficulties are also encountered in the studies of non-uniformly expanding (non-invertible) maps, as it is pointed out in [48]. Recently, some progress has been made in this direction for some hyperbolic maps, see [53, 56].

In this paper we are able to identify the main factors that affect the decay rates of correlations for rather general nonuniformly hyperbolic systems, in terms of the tail distribution of the return time function (used in the inducing scheme). Since the singularities of the systems make it difficult to apply all existing methods, we revisit Dolgopyat's coupling method and the ideas of standard pairs [20, 39] for systems with exponential decay rates of correlations,

see also [72], and develop a new coupling scheme for nonuniformly hyperbolic systems. Combining with the elegant Probability Renewal Theory, originated from Kolmogorov [51], we are able to obtain an optimal bound for the decay rates of correlations for general 2-dimensional hyperbolic systems. Our formulas give a precise asymptotic, rather than upper or lower bounds. To our knowledge, this is the first result of that type in the context of nonuniformly hyperbolic billiards. And the results have greatly improved all existing results on decay rates of correlations for hyperbolic systems with slow decay rates of correlations. Moreover, our new method is more flexible, comparing to all other existing methods in this direction, as it can be applied to dynamical systems under small deterministic or random perturbations. In Section 10, we will describe several classes of billiards to which our results can be applied. Moreover we also obtain the optimal bound for an example of nonuniform link-twist map. Other nonuniform hyperbolic maps, some of which exhibiting attractors, have been investigated in another paper [68].

We describe the structure of our work in more details below.

1.2 Plan of exposition

We first prove an upper bound on correlations under rather general assumptions (compared to those in [28, 31]), and then we obtain optimal estimates for the decay of correlations for certain hyperbolic systems, including semi-dispersing billiards, billiards with cusps, Bunimovich stadia, etc. The key ingredient of the coupling scheme is the construction of a Markov tower, which we will call *generalized Young tower* together with a new version of the coupling lemma for nonuniformly hyperbolic systems. Our general scheme consists of three major steps:

(**a**) We first need to choose a subset $M \subset \mathcal{M}$ on which the induced (first return) map $F\colon M \to M$ is uniformly hyperbolic (with an exponential decay of correlations). We note that F preserves the measure μ conditioned on M, which we denote by μ_M; it will be an SRB measure. For the definition of SRB measure, see Section 2.1.

(**b**) Then we check that standard pairs and standard families (which were introduced in [20, 39, 26]), for the induced system (F, M, μ_M) satisfy the specific conditions of our earlier work [32]. These would imply a Coupling Lemma for the induced system, as well as an exponential decay of correlations. Based on the Coupling Lemma for the induced system [32], a generalized Young tower for the induced system is constructed.

(**c**) In order to prove a Coupling Lemma for the original system $(\mathcal{M}, \mathcal{F}, \mu)$, a generalized Markov tower is carefully constructed based on that for the induced map. Since nonuniform hyperbolicity prevents us from using a standard coupling procedure with an exponentially decaying coupling time [20, 26], we have to develop a different approach: we decompose the orginal measure accord-

ing to a stopping time function and perform coupling only at those stopping times. On the one hand, this procedure allows us to "match" images of our measures efficiently when they both properly return to a reference set. On the other hand, the probability renewal theory enables us to keep track of points that have properly returned to the base of the tower sufficiently many times, but failed to couple for various reasons. Our procedure can be applied to 2D (nonuniformly) hyperbolic systems with or without singularities.

The quantitative part of our scheme involves the following estimates. For the set $M \subset \mathcal{M}$ selected in part (**a**), let $R\colon \mathcal{M} \to \mathbb{N}$ be the first hitting time to M under iterations of $\mathcal{F}$; see precise definition in (3.1). The Poincaré recurrence theorem implies that $R < \infty$ almost everywhere on $\mathcal{M}$. We obtain the following upper bound on the decay of correlations in Theorem 2.2 for piecewise Hölder observables f, g on $\mathcal{M}$:

$$|C_n(f, g, \mathcal{F})| \leq C_{f,g}\, \mu(R > n), \tag{1.3}$$

where $C_{f,g} > 0$ is a constant depending only on f and g. Moreover the following formula is derived in Theorem 2.3 for a large class of systems:

$$\int_{\mathcal{M}} (f \circ \mathcal{F}^n)\, g\, d\mu - \int_{\mathcal{M}} f\, d\mu \int_{\mathcal{M}} g\, d\mu = \mu(R > n)\mu(f)\mu(g) + o(\mu(R > n)), \tag{1.4}$$

under the condition that $\mathrm{supp}(f) \subset M$ and $\mathrm{supp}(g) \subset M$. This is a novel result for billiards. Our asymptotic formula (1.4) implies that the tail bound for the return time function $\mu(R > n)$ is indeed an optimal estimate for the decay of correlations. Note that if one chooses certain observables satisfying conditions in Theorem 2 as well as $\mu(f)\mu(g) = 0$, then we get a faster decay rate of correlations:

$$|C_n(f, g, \mathcal{F})| = o(\mu(R > n)), \tag{1.5}$$

as it is proved in Theorem 2.

2 Assumptions and main results

2.1 Assumptions

Let $\mathcal{M}$ be a 2-dimensional compact Riemannian manifold, possibly with boundary. Let $\Omega \subset \mathcal{M}$ be an open subset and let $\mathcal{F}\colon \Omega \to \mathcal{M}$ be a $C^{1+\gamma_0}$ diffeomorphism of Ω onto $\mathcal{F}(\Omega)$ (here $\gamma_0 \in (0, 1]$). We assume that $\mathcal{S}_1 = \mathcal{M} \setminus \Omega$ is a finite or countable union of smooth compact curves. Similarly, $\mathcal{S}_{-1} = \mathcal{M} \setminus \mathcal{F}(\Omega)$ is a finite or countable union of smooth compact curves. If $\mathcal{M}$ has boundary $\partial\mathcal{M}$, it must be a subset of both $\mathcal{S}_1$ and $\mathcal{S}_{-1}$. We call $\mathcal{S}_1$ and $\mathcal{S}_{-1}$ the *singularity sets* for the maps $\mathcal{F}$ and $\mathcal{F}^{-1}$, respectively. We denote by Ω_i, $i \geq 1$, the connected components of Ω; then $\mathcal{F}(\Omega_i)$ are the connected components of $\mathcal{F}(\Omega)$. We assume that $\mathcal{F}|_\Omega$ is **time-reversible**, and the restriction of the map $\mathcal{F}$ to any component Ω_i can be extended by continuity to its boundary $\partial\Omega_i$, though the extensions to

$\partial\Omega_i \cap \partial\Omega_j$ for $i \neq j$ need not agree. Similarly, for each i the restriction of $\mathcal{F}^{-1}$ to any connected component $\mathcal{F}(\Omega_i)$ can be extended by continuity to its boundary $\partial\mathcal{F}(\Omega_i)$.

Next we assume that the map $\mathcal{F}$ is (nonuniformly) hyperbolic, as defined by Katok and Strelcyn [50]. This means that $\mathcal{F}$ preserves a probability measure μ such that μ-a.e. point $x \in \mathcal{M}$ has two non-zero Lyapunov exponents: one positive and one negative. Also, the first and second derivatives of the maps $\mathcal{F}$ and $\mathcal{F}^{-1}$ do not grow too rapidly near their singularity sets $\mathcal{S}_1$ and $\mathcal{S}_{-1}$, respectively; and the ε-neighborhood of the singularity set has measure $\mathcal{O}(\varepsilon^{q_0})$ for some $q_0 > 0$. This is to ensure the existence and absolute continuity of stable and unstable manifolds at μ-a.e. point. Let

$$\mathcal{W}^u = \cap_{n\geq 0}\mathcal{F}^n(\mathcal{M} \setminus \mathcal{S}_1).$$

Obviously, $\mathcal{W}^u$ is (mod 0) the union of all unstable manifolds, and we assume that the partition $\mathcal{W}^u$ of $\mathcal{M}$ into unstable manifolds is measurable, so that μ induces conditional distributions on μ-almost all unstable manifolds (see the definition and basic properties of conditional measures in [26, Appendix A]). Most importantly, we assume that μ is an Sinai-Ruelle-Bowen (SRB) measure; i.e. the conditional distributions of μ on unstable manifolds $W \subset \mathcal{W}^u$ are absolutely continuous with respect to the Lebesgue measure on W. SRB measures are known to be the only physically observable measures, in the sense that their basins of attraction have positive Lebesgue volume; see [73] and [45, Sect. 5.9]. We also **assume that our SRB measure μ is ergodic and mixing**. Our work is devoted to statistical properties of μ, so it is natural for us to take its ergodicity and mixing for granted.

In chaotic billiards, all the above assumptions are satisfied and are usually easy to check. In particular, the invariant measure μ for billiards is smooth and has a positive density on all of Ω. In physics terms, this invariant measure μ is an equilibrium state. Another important class of systems consists of small perturbations of chaotic billiards (usually induced by external forces or special boundary conditions) [18, 19]. Those systems model electrical current [22, 23], heat conduction and viscous flows [14, 25], the motion under gravitation on the Galton board [21], etc. For perturbed billiards all the above assumptions are satisfied, too, but the measure μ is no longer absolutely continuous: it is singular with respect to the Lebesgue measure on $\mathcal{M}$ (though every open subset $U \subset \mathcal{M}$ still has a positive μ-measure). In physics, such a measure μ (for billiards under small perturbations) is called a nonequilibrium steady state (NESS).

Even with the assumptions made thus far, the decay of correlations may be arbitrarily slow [28, 29]. In order to ensure a specific rate for the decay of correlations we introduce the inducing scheme that were used by Markarian [55] for Bunimovich Stadia, Chernov and Zhang [28] for general hyperbolic systems with slow decay of correlations.

We first construct a subset $M \subset \mathcal{M}$, with $\mu(M) > 0$, and let $R : \mathcal{M} \to \mathbb{N}$ be the first hitting time of M. By the Poincare Recurrence Theory, there exists $\hat{\mathcal{M}} \subset \mathcal{M}$, such that $\mu(\hat{\mathcal{M}}) = 1$ and for any $x \in \hat{\mathcal{M}}$, $R(x) < \infty$. For

any $x \in M \cap \hat{\mathcal{M}}$, we define $Fx = \mathcal{F}^{R(x)}$. Then F preserves a probability measure $\mu_M := \mu|_M/\mu(M)$. For precise construction of the induced system, see Subsection 3.1. We next make more specific assumptions on the induced system (M, F, μ_M).

(**H**1) **Distribution of the first hitting time** R**.** We assume that there exist $C > 0$, $\alpha_0 > 1$, such that the distribution of R satisfies:

$$\mu(M_n) \leq \frac{C}{n^{1+\alpha_0}} \tag{2.1}$$

where M_n is the closure of the level set of R restricted on M which is $\{x \in M : R(x) = n\}$. Moreover, we assume that there exists $N_0 \geq 1$ such that every level set M_n contains at most N_0 connected components.

Remark.

(i) *Note that (2.1) implies that there exist $C_1 > 0$, such that*

$$\mu(R > n) = \mu(M) \sum_{k \geq n} \mu_M(R > k) \leq \frac{C_1}{n^{\alpha_0 - 1}}$$

(ii) *In assumption (**H1**), we can replace (2.1) by the following limit:*

$$\lim_{n \to \infty} L(n) n^{1+\alpha_0} \mu(M_n) = C,$$

where $L(n)$ is a slowly vary function at infinity. Our results on decay rates of correlations still hold by simply adjusting the order of the tail bound $\mu(R > n)$ by a slowly vary function in all estimations.

For a large b (whose precise value will be given in (6.17)), we denote $\psi(n) := (b \ln n)^2$, and define the set

$$C_{n,b} = \{x \in \mathcal{M} \,|\, \#_{1 \leq i \leq n}\{\mathcal{F}^i(x) \in M\} < \psi(n),\ R(x) \leq n\} \tag{2.2}$$

Clearly, $C_{n,b}$ contains those points in $\mathcal{M}$ whose forward trajectory only returns to M at most $\psi(n)$ times within n collisions.

(**H**2) **Measure of** $C_{n,b}$**.** We assume that there exists $\varepsilon_1 \in [0, 1)$ such that

$$\mu(C_{n,b}) \leq C n^{1-\alpha_0}; \quad \mu(C_{n,b} \cap \mathcal{F}^{-n} M) \leq C n^{1-\alpha_0-\varepsilon_1}.$$

Next we introduce sufficient conditions for exponential decay rates of correlations, as well as for the coupling lemma, for the induced map. These assumptions are quite standard and have been made in many references [17, 20, 26, 32].

(**H**3) **Sufficient conditions for exponential decay of correlations of the induced map.**

(**h1**) **Hyperbolicity**[2] **of** F. There exist two families of cones C_x^u (unstable) and C_x^s (stable) in the tangent spaces $\mathcal{T}_x M$, for all $x \in M \setminus S_1$, and there exists a constant $\Lambda > 1$, with the following properties:

(1) $D_x F(C_x^u) \subset C_{Fx}^u$ and $D_x F(C_x^s) \supset C_{Fx}^s$, wherever $D_x F$ exists.

(2) $\|D_x F(v)\| \geq \Lambda \|v\|, \forall v \in C_x^u$; and $\|D_x F^{-1}(v)\| \geq \Lambda \|v\|, \forall v \in C_x^s$.

(3) These families of cones are continuous on M and the angle between C_x^u and C_x^s is uniformly bounded away from zero.

We say that a smooth curve $W \subset M$ is an unstable (stable) *curve* if at every point $x \in W$ the tangent line $\mathcal{T}_x W$ belongs in the unstable (stable) cone C_x^u (C_x^s). Furthermore, a curve $W \subset M$ is an unstable (stable) *manifold* if $F^{-n}(W)$ is an unstable (stable) curve for all $n \geq 0$ (resp. ≤ 0).

(**h2**) **Singularities.** The boundary ∂M is transversal to both stable and unstable cones. Every other smooth curve $W \subset S_1 \setminus \partial M$ (resp. $W \subset S_{-1} \setminus \partial M$) is a stable (resp. unstable) curve. Every curve in S_1 terminates either inside another curve of S_1 or on the boundary ∂M. A similar assumption is made for S_{-1}. Moreover, there exist $q_0 \in (0, 1]$ and $C > 0$ such that for any $x \in M \setminus S_1$

$$\|D_x F\| \leq C \operatorname{dist}(x, S_1)^{-q_0}, \tag{2.3}$$

and for any $\varepsilon > 0$,

$$\mu\big(x \in M \colon \operatorname{dist}(x, S_1) < \varepsilon\big) < C\varepsilon^{q_0}. \tag{2.4}$$

Note that (2.4) implies that for μ-a.e. $x \in M$, there exists a stable and unstable manifold $W^{u/s}(x)$, such that $F^n W^s(x)$ and $F^{-n} W^u(x)$ never hit S_1, for any $n \geq 0$.

Definition 2.1. For every $x, y \in M$, define $\mathbf{s}_+(x, y)$, the forward *separation time* of x, y, to be the smallest integer $n \geq 0$ such that x and y belong to distinct elements of $M \setminus S_n$. Fix $\beta \in (0, 1)$, then $d(x, y) = \beta^{\mathbf{s}_+(x,y)}$ defines a metric on M. Similarly we define the backward separation time $\mathbf{s}_-(x, y)$.

(**h3**) **Regularity of stable/unstable manifolds**. We assume that the following families of stable/unstable curves, denoted by $\mathcal{W}_F^{s,u}$ are invariant under F^{-1} (resp., F) and include all stable/unstable manifolds:

(1) **Bounded curvature.** There exist $B > 0$ and $c_M > 0$, such that the curvature of any $W \in \mathcal{W}_F^{s,u}$ is uniformly bounded from above by B, and the length of the curve $|W| < c_M$.

[2] We have already assumed that Lyapunov exponents are not zero a.e., but our methods also use stable and unstable cones for the map F.

(2) **Distortion bounds.** There exist $\gamma_0 \in (0,1)$ and $C_{\mathbf{r}} > 1$ such that for any unstable curve $W \in \mathcal{W}_F^u$ and any $x, y \in W$,

$$\left|\ln \mathcal{J}_W(F^{-1}x) - \ln \mathcal{J}_W(F^{-1}y)\right| \le C_{\mathbf{r}} \operatorname{dist}(x,y)^{\gamma_0}, \tag{2.5}$$

where

$$\mathcal{J}_W(F^{-1}x) = dm_{F^{-1}W}(F^{-1}x)/dm_W(x)$$

denotes the Jacobian of F^{-1} at $x \in W$ with respect to the Lebesgue measure on the unstable curve W.

(3) **Absolute continuity.** Let $W_1, W_2 \in \mathcal{W}_F^u$ be two unstable curves close to each other. Denote

$$W_i' = \{x \in W_i \colon W^s(x) \cap W_{3-i} \neq \emptyset\}, \qquad i = 1, 2.$$

The map $\mathbf{h}\colon W_1' \to W_2'$ defined by sliding along stable manifolds is called the *holonomy* map. We assume $\mathbf{h}_* m_{W_1'}$ is absolutely continuous with respect to $m_{W_2'}$, i.e. $\mathbf{h}_* m_{W_1'} \prec m_{W_2'}$; and furthermore, there exist uniform constants $C_{\mathbf{r}} > 0$ and $\vartheta_0 \in (0,1)$, such that the Jacobian of $\mathbf{h}$ satisfies

$$|\ln \mathcal{J}\mathbf{h}(y) - \ln \mathcal{J}\mathbf{h}(x)| \le C_{\mathbf{r}} \vartheta_0^{\mathbf{s}_+(x,y)}, \qquad \forall x, y \in W_1' \tag{2.6}$$

Similarly, for any $n \ge 1$ we can define the holonomy map

$$\mathbf{h}_n = F^n \circ \mathbf{h} \circ F^{-n} : F^n W_1 \to F^n W_2,$$

and then (2.6) and the uniform hyperbolicity (**h1**) imply

$$\ln \mathcal{J}\mathbf{h}_n(F^n \mathbf{x}) \le C_{\mathbf{r}} \vartheta_0^n. \tag{2.7}$$

(**h4**) **One-step expansion.** We have

$$\liminf_{\delta \to 0} \sup_{W\colon |W| < \delta} \sum_n \left(\frac{|W|}{|V_n|}\right)^{q_0} \cdot \frac{|F^{-1}V_n|}{|W|} < 1, \tag{2.8}$$

where the supreme is taken over regular unstable curves $W \subset M$, $|W|$ denotes the length of W, and V_n, $n \ge 1$, denote the smooth components of $F(W)$.

Note that the boundary $\partial\mathcal{D}$ is a part of the singular set S_1, hence every stable manifold for F is also a stable manifold for $\mathcal{F}$. Since we denote by $\mathcal{W}_F^s$ the collection of all stable manifolds for F, then the collection $\mathcal{W}^s$ of stable manifolds for $\mathcal{F}$ can be constructed by

$$\mathcal{W}_*^s = \cup_{m=1}^{\infty} \cup_{k=0}^{m-1} \mathcal{F}^k \{W^s \in \mathcal{W}_F^s \colon W^s \subset \mathcal{D}_m\}. \tag{2.9}$$

On the other hand, every unstable manifold W_F^u for F is a (part of) an unstable manifold W^u for $\mathcal{F}$, more precisely, $W_F^u = W^u \cap \Omega_i$ for some $\Omega_i \subset \mathcal{D}$. Since

we denote by $\mathcal{W}_F^u$ the collection of all unstable manifolds for F, then it can be extended to the whole space $\mathcal{M}$ in a similar way:

$$\mathcal{W}_*^u = \cup_{m=1}^{\infty} \cup_{k=0}^{m-1} \mathcal{F}^{-k}\{W^u \in \mathcal{W}_F^u \colon W^u \subset F(\mathcal{D}_m)\}. \tag{2.10}$$

Notice that this will not be exactly the collection $\mathcal{W}^u$ of unstable manifolds for $\mathcal{F}$, instead the latter would be obtained by concatenation of some curves from $\mathcal{W}_*^u$. Since the collection of curves in $\mathcal{W}^{u/s} \setminus \mathcal{W}_*^{u/s}$ is a null set, we will not make distinction between these two sets below. More precisely, we will identify $\mathcal{W}^\sigma = \mathcal{W}_*^\sigma$, for $\sigma \in \{u, s\}$.

2.2 Statement of the main results

For any $\gamma \geq \gamma_0$, let $\mathcal{H}^-(\gamma)$ be the set of all bounded real-valued functions $f \in L_\infty(\mathcal{M}, \mu)$ such that for any x and y lying on one stable manifold $W^s \in \mathcal{W}^s$,

$$|f(x) - f(y)| \leq \|f\|_\gamma^- \mathrm{dist}(x, y)^\gamma, \tag{2.11}$$

with

$$\|f\|_\gamma^- \colon = \sup_{W^s \in \mathcal{W}^s} \sup_{x,y \in W^s} \frac{|f(x) - f(y)|}{\mathrm{dist}(x, y)^\gamma} < \infty.$$

Similarly, we define $\mathcal{H}^+(\gamma)$ as the set of all bounded, real-valued functions $g \in L_\infty(\mathcal{M}, \mu)$ such that for any x and y lying on one unstable manifold $W^u \in \mathcal{W}^u$,

$$|g(x) - g(y)| \leq \|g\|_\gamma^+ \mathrm{dist}(x, y)^\gamma, \tag{2.12}$$

with

$$\|g\|_\gamma^+ \colon = \sup_{W^u \in \mathcal{W}^u} \sup_{x,y \in W^u} \frac{|g(x) - g(y)|}{\mathrm{dist}(x, y)^\gamma} < \infty.$$

When we study autocorrelations of certain observables, we will need to require that the latter belongs to the space $\mathcal{H}(\gamma) \colon = \mathcal{H}^+(\gamma) \cap \mathcal{H}^-(\gamma)$, i.e., they are Hölder functions on every stable and unstable manifold. For correlations of two distinct functions given by (1.1) we always assume that $f \in \mathcal{H}^-(\gamma_1)$ and $g \in \mathcal{H}^+(\gamma_2)$ with some $\gamma_1, \gamma_2 \in [\gamma_0, 1]$, unless otherwise specified. For every $f \in \mathcal{H}^\pm(\gamma)$ we define

$$\|f\|_{C^\gamma}^\pm \colon = \|f\|_\infty + \|f\|_\gamma^\pm. \tag{2.13}$$

By using the coupling methods, we obtain the following upper bounds for the rate of decay of correlations.

Theorem 2.2. *For systems satisfy (**H1**)-(**H3**), there exists $C > 0$, such that for any observables $f \in \mathcal{H}^-(\gamma_1)$ and $g \in \mathcal{H}^+(\gamma_2)$ on $\mathcal{M}$, with $\gamma_i \in [\gamma_0, 1]$, for $i = 1, 2$,*

$$\begin{aligned} |\mu(f \circ \mathcal{F}^n \cdot g) - \mu(f)\mu(g)| &\leq C\|g\|_{C^{\gamma_2}}^+ \|f\|_{C^{\gamma_1}}^- (\mu(C_{n,b}) + \mu(R > n)) \\ &\leq C\|g\|_{C^{\gamma_2}}^+ \|f\|_{C^{\gamma_1}}^- n^{1-\alpha_0}, \end{aligned}$$

*for $n \geq N$, with $\alpha_0 > 1$ was given in (**H1**) and $N_1 = N_1(g) \geq 1$.*

Note that (**H1**) implies that $\mu(R > n) = \mathcal{O}(n^{1-\alpha_0})$. Thus the decay rates of correlations are closely related to the tail distribution of the function R. Next we indeed show that $\mu(R > n)$ characterizes the optimal bound for the decay rates of correlations for a large class of observables.

Theorem 2.3. *Under conditions of Theorem1, if we further assume both f and g supported in M (note M is the nice subset in $\mathcal{M}$), then correlations decay as:*

$$\mu(f \circ \mathcal{F}^n \cdot g) - \mu(f)\mu(g) = \mu(R > n)\mu(f)\mu(g)| + E(f, g, n), \tag{2.14}$$

for any $n \geq N_1$, with

$$|E(f, g, n)| \leq C\|f\|_\infty \|g\|_\infty n^{-\beta_0} = o(\mu(R > n)),$$

*and $\beta_0 = \min\{\alpha_0 + \varepsilon_1 - 1, 2\alpha_0 - 2\}$. Here $\varepsilon_1 \in [0, 1)$ was given in (**H2**).*

For the case where $\mu(f)\mu(g) = 0$ we also have a better bound $o(\mu(R > n))$ specified in the above Theorem 2. For example, for semi-dispersing billiards on a rectangle and Bunimovich Stadia, we have $\alpha_0 = 2$, so for general observables, the correlations decay as $O(n^{-1})$; see also [55, 28]. But for observables supported on M and such that $\mu(f)\mu(g) = 0$, the above theorem implies $C_n(f, g, \mathcal{F}) = O(n^{-1-\varepsilon_1})$, which also implies the classical central limit theorem.

Next we will prove that for dynamical systems with slow decay rates of correlations, the class of Hölder observables f with support on M will satisfy the classical Central Limit Theorem.

Theorem 2.4. *Assume $\beta_0 > 1$ in Theorem 2. Let $f \in \mathcal{H}^-(\gamma_0) \cap \mathcal{H}^+(\gamma_0)$ with $\text{supp}(f) \subset M$ and $\mu(f) = 0$. Assume f is not a coboundary, i.e. there is no function h such that $f = h - h \circ \mathcal{F}$. Then the following sequence converges:*

$$\frac{f + \cdots + f \circ \mathcal{F}^n}{\sigma\sqrt{n}} \xrightarrow{d} Z, \tag{2.15}$$

in distribution, as $n \to \infty$. Here

$$\sigma^2 = \mu(f^2) + 2\sum_{n=0}^{\infty} \mu(f \circ \mathcal{F}^n \cdot f) < \infty$$

and Z is a standard normal variable.

According to Theorem 3, for Bunimovich Stadia and billiards with cusps, even though the correlation for general variables usually decay at order $\mathcal{O}(n^{-1})$, if we pick a Hölder continuous function f supported on M, with $f \in H^+(\gamma_0) \cap H^-(\gamma_0)$ and $\mu(f) = 0$, then we still be able to get a classical Central limit theorem, instead of the abnormal Central limit theorem.

Another related observable we would like to discuss is $\hat{f} := f - \mu(f)$, the centralized version of f, with $\text{supp}(f) \subset M$. Note that we can express $\mu(f) = \mu(f)\mathbf{I}_{\mathcal{M}}$. Thus $\hat{f} = f - \mu(f)\mathbf{I}_{\mathcal{M}}$ is not supported on M any more, so our Theorem 2 does not apply to it. Moreover, to study the Central limit theorem for $\hat{f} \circ \mathcal{F}^n$,

we need to define $\tilde{f} := (f - \mu(f)R) \cdot \mathbf{I}_M$, which is the induced observable by $\hat{f}$. Thus the partial sums

$$\tilde{S}_n := \hat{f} + \hat{f} \circ \mathcal{F} + \cdots + \hat{f} \circ \mathcal{F}^n,$$

and

$$S_n = \tilde{f} + \tilde{f} \circ F + \cdots + \tilde{f} \circ F,$$

should have similar asymptotic behavior if we scale them by the reciprocal of their standard deviation. However, by the definition of $\tilde{f}$, we know that the variance of $\tilde{f}$ is dominated from below by that of R, which is unbounded for the case when $\mu(R > n) \sim n^{-1}$. This means that the classical Central limit fails, as well as the Green-Kubo formula, as

$$\sum_{n=1}^{\infty} \mu(R > n) \sim \sum_{n=1}^{\infty} n^{-1} = \infty.$$

For the Central Limit Theorem of these general classes of observable for Bunimovich Stadia and billiards with cusps, see [6, 5].

Throughout the paper we will use the following conventions: Positive and finite global constants whose value is unimportant, will be denoted by $c, c_1, c_2, \cdots$ or $C, C_1, C_2, \cdots$. These letters may denote different values in different equations throughout the paper. In Appendix, we also list some notations that we use throughout the paper.

3 Standard families and the induced map

In this paper, we will use the coupling method to prove our main theorems, which depends heavily on the concept of standard pairs proposed firstly by Dolgopyat in [39], as well as the $\mathcal{Z}$ function by Chernov and Dolgopyat in [20, 26].

3.1 Construction of an induced system (F, M, μ_M).

In this section, we carefully construct the induced system, and take care of the definition of singularity set for both systems.

Let $\mathcal{M}$ be a 2-dimensional compact Riemannian manifold, and $\Omega \subset \mathcal{M}$ be an open subset. $\mathcal{S}_1 = \mathcal{M} \setminus \Omega$ and $\mathcal{S}_{-1} = \mathcal{M} \setminus \mathcal{F}(\Omega)$ are the *singularity sets* for the maps $\mathcal{F}$ and $\mathcal{F}^{-1}$, respectively. Let Ω_i, $i \geq 1$, be the connected components of Ω; then $\mathcal{F}(\Omega_i)$ are the connected components of $\mathcal{F}(\Omega)$.

Let $\mathcal{D} = \cup_{i \in I} \Omega_i$ ($\operatorname{card} I < \infty$) be a finite union of some connected components of Ω. For any $x \in \mathcal{M}$, let

$$R(x) = \min\{n \geq 1 \colon \mathcal{F}^n(x) \in \mathcal{D}, \ \ \mathcal{F}^m(x) \notin \mathcal{S}_1, \ m = 1, \ldots, n-1\}, \tag{3.1}$$

be the first entrance time to the set $\mathcal{D}$. When restricted on $\mathcal{D}$, we also call R the first return time, although later we extend it to almost every points of $\mathcal{M}$.

We denote by $\mathcal{N}_1 \subset \mathcal{D}$ the set of points that never return to $\mathcal{D}$; it consists of points of two types:

(i) $\mathcal{F}^n(x) \in \mathcal{S}_1$ for some $n \geq 1$ and $\mathcal{F}^m(x) \notin \mathcal{D}$ for $m = 1, \ldots, n-1$ (the orbit of x hits a singularity before it comes back to $\mathcal{D}$);

(ii) $\mathcal{F}^n(x) \in \Omega \setminus \mathcal{D}$ for all $n \geq 1$ (i.e., x is a wandering point).

For each $n \geq 1$, the "level" set

$$\mathcal{D}_n := \{x \in \mathcal{D} \,:\, R(x) = n\} \subset \mathcal{D}$$

is open, and if $\mathcal{D}_n \neq \emptyset$ then $\mathcal{F}^n$ is a diffeomorphism of $\mathcal{D}_n$ onto $\mathcal{F}^n(\mathcal{D}_n) \subset \mathcal{D}$. We denote by F the first return map, i.e.,

$$F(x) = \mathcal{F}^n(x) \qquad \forall x \in \mathcal{D}_n, \quad n \geq 1.$$

It is easy to see that F is a diffeomorphism of the open set $\mathcal{D}^+ = \cup_{n\geq 1}\mathcal{D}_n$ onto the open set $\mathcal{D}^- = \cup_{n\geq 1}\mathcal{F}^n(\mathcal{D}_n)$. The inverse map F^{-1} is defined on $\mathcal{D}^- \subset \mathcal{D}$ and takes it back to $\mathcal{D}^+$. Let $M = \overline{\mathcal{D}}$ denote the closure of $\mathcal{D}$, and for each $n \geq 1$ let $M_n = \overline{\mathcal{D}_n}$. We put

$$S_1 = M \setminus \mathcal{D}^+ = \mathcal{N}_1 \cap \partial\mathcal{D}$$

and

$$S_{-1} = M \setminus \mathcal{D}^- = \mathcal{N}_{-1} \cap \partial\mathcal{D},$$

where $\mathcal{N}_{-1} \subset \mathcal{D}$ denotes the set of points never coming back to $\mathcal{D}$ under the iterations of $\mathcal{F}^{-1}$. We assume that both S_1 and S_{-1} are finite or countable unions of smooth compact curves. The sets $S_{\pm 1}$ play the role of singularities for the induced maps $F^{\pm 1}$. We assume that the map F restricted to any level set $\mathcal{D}_n$ can be extended by continuity to its boundary $\partial\mathcal{D}_n$, but the extensions to $\partial\mathcal{D}_n \cap \partial\mathcal{D}_m$ for $n \neq m$ need not agree. A similar assumption is made for F^{-1}.

We assume that $\mu(\mathcal{D}) > 0$. It is easy to show [24] that the SRB measure μ cannot be concentrated on curves, i.e., $\mu(W) = 0$ for any smooth curve $W \subset \mathcal{M}$. Thus all our singularity sets $\mathcal{S}_{\pm 1}$, $S_{\pm 1}$, and their images under $\mathcal{F}^n$, $n \in \mathbb{Z}$, are null sets. By the ergodicity of μ we have

$$\mathcal{M} = \bigcup_{n\geq 1}\bigcup_{m=0}^{n-1} \mathcal{F}^m\mathcal{D}_n = \bigcup_{m\geq 0}\bigcup_{n=m+1}^{\infty} \mathcal{F}^m\mathcal{D}_n = \bigcup_{m\geq 1} R^{-1}\{m\} \qquad (\text{mod } 0), \quad (3.2)$$

where

$$R^{-1}\{m\} = \bigcup_{n=m}^{\infty} \mathcal{F}^{n-m}\mathcal{D}_n$$

is the level set of R in $\mathcal{M}$. We note that $\int_M R\, d\mu = 1$ by the Kac theorem. The first return map F preserves the measure μ conditioned on M; we denote it by μ_M. Clearly, this measure is ergodic. We also **assume that μ_M is mixing**.

Unstable manifolds $W \subset M$ are the unstable manifolds for the induced map F intersected with M, hence μ_M is an SRB measure.

For $n \geq 1$, let

$$S_n = \bigcup_{i=0}^{n-1} F^{-i} S_1 \quad \text{and} \quad S_{-n} = \bigcup_{i=0}^{n-1} F^i S_{-1},$$

for each $n \geq 1$. Then the map $F^n \colon M \backslash S_n \to M \backslash S_{-n}$ is a $C^{1+\gamma_0}$ diffeomorphism.

3.2 Standard families

For any unstable manifold $W \in \mathcal{W}^u$, let μ_W be the probability measure on W, obtained by conditioning μ on W. More precisely, the probability measure μ_W is determined by the unique probability density ρ_W (with respect to the Lebesgue measure m_W) satisfying

$$\frac{\rho_W(y)}{\rho_W(x)} = \lim_{n\to\infty} \frac{\mathcal{J}_W(F^{-n}y)}{\mathcal{J}_W(F^{-n}x)}. \tag{3.3}$$

Here ρ_W is called *the u-SRB density on* W, and the corresponding probability measure μ_W on W is called *the u-SRB measure of* F. The formula (3.3) is standard in ergodic theory, see [26] page 105.

Definition 3.1 (Standard pair)**.** A probability measure ν supported on an unstable manifold W is called *regular*, if ν is absolutely continuous with respect to the u-SRB measure μ_W, and the probability density function $g = d\nu/d\mu_W$ satisfies

$$|\ln g(x) - \ln g(y)| \leq C_F d(x,y)^{\gamma_0}, \tag{3.4}$$

where $\gamma_0 \in (0,1)$ was given in (2.5), and $C_F > C_{\mathbf{r}}$ is a fixed large constant. In this case (W,ν) is called a *standard pair*. Moreover, if the probability density $g = d\nu/d\mu$ satisfies

$$|\ln g(x) - \ln g(y)| \leq 4C_F d(x,y)^{\gamma_0}, \tag{3.5}$$

we call (W,ν) a pseudo-standard pair.

We need to introduce the concept of pseudo-standard pair, as in the proof of the Coupling Lemma, one needs to subtract a smooth function from the density of a standard pair. But the resulting conditional measure will only induce a pseudo-standard pair, as we will show in lemma 4.3.

Definition 3.2. Let $\mathcal{G} = \{(W_\alpha, \nu_\alpha), \alpha \in \mathcal{A}\}$ be a family of standard pairs equipped with a factor measure λ on the index set $\mathcal{A}$. We call $\mathcal{G}$ a *standard family* on $\mathcal{M}$, if it satisfies the following conditions:

(i) $\{W_\alpha, \alpha \in \mathcal{A}\}$ is a measurable partition of a measurable subset $\mathcal{W} \subset \mathcal{M}$ into unstable manifolds;

(ii) There is a Borel probability measure ν satisfying:

$$\nu(B) = \int_{\alpha\in\mathcal{A}} \nu_\alpha(B \cap W_\alpha)\, \lambda(d\alpha), \tag{3.6}$$

for any measurable set $B \subset \mathcal{M}$.

For simplicity, we denote such a family by

$$\mathcal{G} = (\mathcal{W}, \nu) = \{(W_\alpha, \nu_\alpha), \alpha \in \mathcal{A}, \lambda\}.$$

We denote $\mathcal{H}(\mathcal{M})$ as the space of bounded, Hölder functions, and for any $\varphi \in \mathcal{H}(\mathcal{M})$, we define

$$\mathcal{G}(\varphi) := \int_{\alpha\in\mathcal{A}} \int_{x\in W_\alpha} \varphi(x)\, d\nu_\alpha\, \lambda(d\alpha). \tag{3.7}$$

Then one can check that any standard family $\mathcal{G}$ on $\mathcal{M}$ is a bounded linear functional (or generalized function) on the space of test functions $\mathcal{H}(\mathcal{M})$. We denote $\mathfrak{F}(\mathcal{M})$ ($\mathfrak{F}(M)$) as the collection of all standard families on $\mathcal{M}$ (resp. M). Both sets are closed under positive scalar multiplications.

Moreover, since a standard family can be viewed as a weighted sum of standard pairs, then for any collection of standard families $\mathcal{G}_i = (\mathcal{W}_i, \nu_i) \in \mathfrak{F}(\mathcal{M})$, $i \geq 1$, the following sum is well-defined:

$$\sum_{i=1}^{\infty} \mathcal{G}_i := \left(\bigcup_{i=1}^{\infty} \mathcal{W}_i, \sum_{i=1}^{\infty} \nu_i \right), \tag{3.8}$$

It is also a standard family in $\mathfrak{F}(\mathcal{M})$, as long as $\{\mathcal{W}_1, \cdots, \mathcal{W}_n, \cdots\}$ are mutually disjoint, and $\sum_{i=1}^{\infty} \nu_i(\mathcal{M}) < \infty$.

It was shown in [39, 26, 32] that $\mathfrak{F}(M)$ is invariant under F, and $(\mathcal{W}_F^u, \mu_M) \in \mathfrak{F}(M)$.

More precisely, if $\mathcal{G} = (\mathcal{W}, \nu)$ is a standard family with a factor measure λ then $F_*^n \nu$ induces a standard family with $F^n\mathcal{G} = F^n(\mathcal{W}, \nu) := (F^n\mathcal{W}, F_*^n\nu)$. Here, we denote

$$F^n(\mathcal{W}, \nu) = ((V_\alpha, \nu_\alpha) \,:\, \alpha \in \mathcal{A}^n, \lambda_n) \tag{3.9}$$

as the standard family with factor measure λ_n on the index set $\mathcal{A}^n$ of unstable manifolds in $F^n\mathcal{W}$.

Moreover, the set $\mathcal{W}_F^u = \{W_\alpha,\, \alpha \in \mathcal{A}_M^u\}$ can be viewed as the measurable partition of M into unstable manifolds. It was proved in [32] Lemma 12 that $(\mathcal{W}_F^u, \mu_M)$ can be viewed as a standard family. Let $\mathcal{A}_M^u$ be the index set such that

$$\mathcal{W}_F^u = \{W_\alpha \,:\, \alpha \in \mathcal{A}_M^u\}.$$

Then μ_M induces a factor measure λ_M^u on the sigma algebra (induced by the Borel sigma algebra on M) of the index set $\mathcal{A}_M^u$, such that

$$(\mathcal{W}_F^u, \mu_M) = \{(W_\alpha, \mu_\alpha), \alpha \in \mathcal{A}_M^u, \lambda_M^u\}. \tag{3.10}$$

In this paper, we always take the u-SRB measure $\mu_\alpha := \mu_{W_\alpha}$ to be the reference measure on indexed unstable manifold $W_\alpha \in \mathcal{W}^u$, sometime we also denote it as μ_W for general unstable manifold $W \in \mathcal{W}^u$.

In particular, for any $m \geq 1$, we define the index set $\mathcal{A}_m^u$ as all unstable manifolds of F in the level set $\mathcal{D}_m$. More precisely,

$$\{W_\alpha \,:\, \alpha \in \mathcal{A}_m^u\} = \{W_\alpha \subset \mathcal{D}_m \,:\, \alpha \in \mathcal{A}_M^u\}. \tag{3.11}$$

One can check that

$$\lambda_M^u(\mathcal{A}_m) = \mu_M(\mathcal{D}_m) = \mu(\mathcal{D}_m)/\mu(M).$$

We denote $(\mathcal{W}^u, \mu)|_{\mathcal{D}_m}$ as the standard family obtained by restricting $(\mathcal{W}^u, \mu)$ on $\mathcal{D}_m$. We will use it to show that $\mathfrak{F}(\mathcal{M})$ is also invariant under $\mathcal{F}$.

Proposition 3.3. For any $\mathcal{G} \in \mathfrak{F}(\mathcal{M})$, then $\mathcal{F}\mathcal{G} \in \mathfrak{F}(\mathcal{M})$. Moreover, $(\mathcal{W}^u, \mu) \in \mathfrak{F}(\mathcal{M})$.

To prove this proposition, we first need the following lemma.

Lemma 3.4. *For any $m \geq 1$, let (W, ν) be a standard pair, with $W \subset \mathcal{D}_m$. Then $\mathcal{F}^k(W, \nu)$ is a standard family, for any $k = 1, \cdots, m$. In particular, $\mathcal{F}^k(W, \nu)$ is a standard pair for $k = 1, \cdots, m-1$.*

Proof. For any $m \geq 1$, let (W, ν) be any standard pair, with $W \in \mathcal{W}_F^u \cap \mathcal{D}_m$. For $k = m$, we know that $\mathcal{F}^m W = FW$. Since F is uniformly hyperbolic, it follows from standard arguments that $\mathcal{F}^m(W, \nu) = F(W, \nu)$ is a standard family. Next we consider the case when $k = 1, \cdots, m-1$. By the definition of $\mathcal{D}_m$, we know that $\mathcal{F}^k W$ is one smooth curve for any $k = 1, \cdots, m-1$.

For the u-SRB measure on W, we denote $\rho_W = d\mu_W/dm_W$. Then (3.3) implies that

$$\begin{aligned}\rho_{\mathcal{F}^k W}(x) &= \frac{d\mathcal{F}_*^k \mu_W(x)}{dm_{\mathcal{F}^k W}(x)} = \frac{d\mu_W(\mathcal{F}^{-k}x)}{dm_{\mathcal{F}^k W}(x)} \\ &= \frac{d\mu_W(\mathcal{F}^{-k}x)}{dm_W(\mathcal{F}^{-k}x)} \cdot \frac{dm_W(\mathcal{F}^{-k}x)}{dm_{\mathcal{F}^k W}(x)} = \frac{\rho_W(\mathcal{F}^{-k}x)}{\mathcal{J}_{\mathcal{F}^k W}(\mathcal{F}^{-k}x)}.\end{aligned}$$

Note that $\mathcal{F}^k W$ contains only one smooth component.

We denote $g = d\nu/d\mu_W$, then for any $k \geq 1$, the density function g_k of $\mathcal{F}_*^k \nu$ can be written as

$$\begin{aligned}g_k(x) &= \frac{d\mathcal{F}_*^k \nu(x)}{d\mu_{\mathcal{F}^k W}(x)} = \frac{d\nu(\mathcal{F}^{-k}x)}{d\mu_W(\mathcal{F}^{-k}x)} \cdot \frac{d\mu_W(\mathcal{F}^{-k}x)}{dm_W(\mathcal{F}^{-k}x)} \cdot \frac{dm_W(\mathcal{F}^{-k}x)}{dm_{\mathcal{F}^k W}(x)} \cdot \frac{dm_{\mathcal{F}^k W}(x)}{d\mu_{\mathcal{F}^k W}(x)} \\ &= \frac{g(\mathcal{F}^{-k}(x))}{\mathcal{J}_{\mathcal{F}^k W}(\mathcal{F}^{-k}(x))} \cdot \frac{\rho_W(\mathcal{F}^{-k}x)}{\rho_{\mathcal{F}^k W}(x)} = g(\mathcal{F}^{-k}(x)),\end{aligned}$$

for all $x \in \mathcal{F}^k W$.

Thus for any x, y belong to one smooth component of $\mathcal{F}^k(W)$, $k \geq 1$,

$$|\ln g_k(x) - \ln g_k(y)| \leq |\ln g(\mathcal{F}^{-k}x) - \ln g(\mathcal{F}^{-k}y)| \leq C_F d(\mathcal{F}^{-k}x, \mathcal{F}^{-k}y)^{\gamma_0}.$$

This implies $(\mathcal{F}^k W, \mathcal{F}^k_* \nu)$ is a standard pair. Moreover, we can check that for any x, y belong to one smooth component of $FW = \mathcal{F}^m(W)$,

$$|\ln g_k(x) - \ln g_k(y)| \leq |\ln g(F^{-1}x) - \ln g(F^{-1}y)| \leq C_F \Lambda^{-\gamma_0} d(x,y)^{\gamma_0}.$$

□

Next we return to the proof of Proposition 3.4. For any standard family $\mathcal{G} = (\mathcal{W}, \nu)$, since the $\mathcal{F}$ image of every standard pair in $\mathcal{G}$ is a standard family, thus $\mathcal{F}\mathcal{G} \in \mathfrak{F}(\mathcal{M})$. Now it remains to show that $(\mathcal{W}^u, \mu) \in \mathfrak{F}(\mathcal{M})$. Let $\mathcal{G} = (\mathcal{W}^u, \mu) = \{(W_\alpha, \mu_\alpha), \alpha \in \mathcal{A}^u, \lambda^u\}$. It follows from the invariance of μ, we have

$$\mu = \sum_{m=1}^{\infty} \sum_{k=1}^{m} \mathcal{F}^k(\mu|_{\mathcal{D}_m})$$

Combining with (2.10), $(\mathcal{W}, \mu)$ can be decomposed as:

$$(\mathcal{W}^u, \mu) = \sum_{m=1}^{\infty} \sum_{k=1}^{m} \mathcal{F}^k((\mathcal{W}^u, \mu)|_{\mathcal{D}_m}) = \sum_{m=1}^{\infty} \mathcal{G}_m.$$

where

$$\mathcal{G}_m := \left(\bigcup_{k=0}^{m-1} \mathcal{F}^k W_{m,\alpha}, \sum_{k=0}^{m-1} \mathcal{F}^k_* \mu_{m,\alpha} \right) = \sum_{k=0}^{m-1} \{\mathcal{F}^k(W_{m,\alpha}, \mu_{m,\alpha}), \alpha \in \mathcal{A}_m, \lambda^u\}$$

Lemma 3.4 implies that for each $m \geq 0$, $\mathcal{G}_m$ is a standard family. Since $\mathcal{G}_m$ all have disjoint support $\cup_{k=0}^{m-1} \mathcal{F}^k(\mathcal{W} \cap \mathcal{D}_m)$, thus by (3.8), we know that $(\mathcal{W}^u, \mu)$ is also a standard family. This finishes the proof of Proposition 3.3. □

To understand the distribution of short unstable manifolds in any standard family, we define a function $\mathcal{Z}$ on standard families. For any standard family $\mathcal{G} = (\mathcal{W}, \nu)$,

$$\mathcal{Z}(\mathcal{G}) = \frac{1}{\nu(\mathcal{M})} \int_{\mathcal{A}} |W_\alpha|^{-1} \, d\lambda(\alpha). \tag{3.12}$$

We fix a large number $C_{\mathrm{q}} > 100 C_F$. Given a standard family $\mathcal{G}$, if $\mathcal{Z}(\mathcal{G}) < C_{\mathrm{q}}$ we say $\mathcal{G}$ is a *proper* family.

Moreover, we choose δ_0 small enough to satisfy

$$\frac{1}{C_{\mathrm{q}}} < \delta_0^{\gamma_0} < \frac{0.01}{C_F}. \tag{3.13}$$

Combining with (3.4), if one makes a standard pair (W, ν) with density function $g \in \mathcal{H}^+(\gamma_0)$ with respect to the conditional measure of μ_W on W, then (3.13) implies that for any $x, y \in W$,

$$|\ln g(x) - \ln g(y)| \leq C_F (20\delta_0)^{\gamma_0} < 0.01. \tag{3.14}$$

Since $g(x) = d\nu/d\mu_W$ is a probability density, there exists $y \in W$, such that $g(y) = 1$. This implies that

$$|g(x) - 1| \leq 0.01. \tag{3.15}$$

Thus for any standard pair with length $|W| < 20\delta_0$, its density function is bounded from below by 0.99, we call it *a proper standard pair*. In particular, the following facts were proved in [32].

Lemma 3.5. *The following statements hold:*
(1) There exists a uniform constant $\chi > 0$, such that for any standard pair (W, ν), $F^n(W, \nu)$ is a proper family for any $n > \chi \ln |W|$;
(2) Let $\mathcal{G} = (\mathcal{W}, \nu)$ be a standard family, then there exists $N > 1$, such that $F^N\mathcal{G}$ is a proper family;
(3) For any $x \in M$, let $\mathbf{r}^{s/u}(x)$ be the distance of x to the boundary of $W^{s/u}(x)$ measured along $W^{s/u}(x)$. Then any standard pair (W, ν) with length $|W| > \delta_0$ is proper; and there exists $C > 0$ such that for any stable/unstable manifold $W^{s/u}$ with length $|W^{s/u}| > \delta_0$, we have

$$m_{W^s}(\mathbf{r}^u(x) < \varepsilon) < C\varepsilon^{q_0}, \quad m_{W^u}(\mathbf{r}^s(x) < \varepsilon) < C\varepsilon^{q_0}; \tag{3.16}$$

(4) There exists $C > 0$, such that for any standard family $\mathcal{G} = ((W_\alpha, \alpha \in \mathcal{A}), \nu)$, any $\varepsilon \in (0, 1)$,

$$\nu(x \in W_\alpha \,:\, \mathbf{r}^u(x) < \varepsilon, \alpha \in \mathcal{A}) < C\mathcal{Z}(\mathcal{G})\varepsilon^{q_0}; \tag{3.17}$$

(5) There exist constants $c > 0$, $C_z > 0$, and $\vartheta_3 \in (0, 1)$, such that for any standard family $\mathcal{G} = (\mathcal{W}, \nu)$ supported in M, any $n \geq 1$,

$$\mathcal{Z}(F^n\mathcal{G}) \leq c\vartheta_3^n\mathcal{Z}(\mathcal{G}) + C_z; \tag{3.18}$$

$$F_*^n\nu(r^{u/s} < \varepsilon) \leq c\,\vartheta_3^n\nu(r^{u/s} < \varepsilon) + C_z\varepsilon^{q_0}. \tag{3.19}$$

Remark. Note that the first equation in (3.16) is the time-reversal version of the second equation, which follows from our assumption **(h2)** on dynamics near singularities.

Definition 3.6. For any standard family $\mathcal{G} = \{(W_\alpha, \nu_\alpha), \alpha \in \mathcal{A}, \lambda\}$, any $n \geq 0$, let $F^n\mathcal{G}$ be defined as in (3.9) with the factor measure λ_n. We define a new standard family using the same factor measure λ_n, but changing the probability measure from ν_α to the u-SRB measure μ_α on each unstable curve W_α:

$$\mathcal{G}_n^u := (F^nW, \mu_n) := ((V_\alpha, \mu_\alpha) \,:\, \alpha \in \mathcal{A}_n, \lambda_n). \tag{3.20}$$

We call $\mathcal{G}_n^u$ the *associated μ-standard family of $F^n\mathcal{G}$*.

Next lemma discusses the relations between these two standard families, which implies that for n large, essentially only the factor measure λ_n matters.

Lemma 3.7. *There exists $C > 0$ such that for any standard family $\mathcal{G} := (\mathcal{W}, \nu)$ with $\mathcal{G}_n^u = (F^n\mathcal{W}, \mu_n)$ being the associated μ-standard family of $F^n\mathcal{G}$, for any $n \geq 1$, then*

$$|F_*^n\nu(A) - \mu_n(A)| \leq C\nu(F^{-n}A)\Lambda^{-\gamma_0 n}, \tag{3.21}$$

for any measurable collection of stable manifolds A.

Proof. We first consider the case when $\mathcal{G} = (W, \nu)$ is a standard pair with density

$$g := d\nu/d\mu_W \in \mathcal{H}^+(\gamma_0).$$

Let

$$F^nW = \{V_\alpha, \alpha \in \mathcal{A}_n\},$$

then $\mathcal{A}_n$ is at most countable. By the definition (3.4), we know that for any $x, y \in V_\alpha$:

$$|\ln g(F^{-n}x) - \ln g(F^{-n}y)| \leq C_F d(F^{-n}x, F^{-n}y)^\gamma \leq C_F|F^{-n}V_\alpha|^{\gamma_0}. \tag{3.22}$$

Or equivalently,

$$|g(F^{-n}x) - g(F^{-n}y)| \leq C_F|F^{-n}V_\alpha|^{\gamma_0} g(F^{-n}y). \tag{3.23}$$

Since μ is invariant, one can check that the push forward measure $F_*^n\nu$ has density $g(F^{-n}x)$, with respect to the SRB measure μ. More precisely, for any measurable collection of stable manifolds $A \subset \mathcal{M}$,

$$\begin{aligned}
F_*^n\nu(A) &= \int \chi_A(x)\, dF^n\nu(x) \\
&= \int \chi_A(x) \cdot \frac{d\nu(F^{-n}x)}{d\mu(F^{-n}x)} \cdot \frac{d\mu(F^{-n}x)}{d\mu(x)}\, d\mu(x) \\
&= \int \chi_A(x) \cdot g(F^{-n}x)\, d\mu(x) \\
&= \sum_{\alpha\in\mathcal{A}_n} \int_{V_\alpha} \chi_A(x) \cdot g(F^{-n}x)\, d\mu_\alpha\, \lambda_{\mathcal{M}}^u(\alpha) \\
&= \sum_{\alpha\in\mathcal{A}_n} \int_{V_\alpha} \chi_A(x) d\nu_\alpha\, \lambda_n(\alpha),
\end{aligned} \tag{3.24}$$

where $d\nu_\alpha = g_\alpha\, d\mu_\alpha$ and

$$\lambda_n(d\alpha) = (\int_{V_\alpha} g(F^{-n}x)\, d\mu_\alpha(x))\lambda_{\mathcal{M}}^u(d\alpha).$$

Here g_α is the density function defined only on V_α such that for any $x \in V_\alpha$,

$$g_\alpha(x) = \frac{g(F^{-n}x)}{\int_{V_\alpha} g(F^{-n}y) d\mu_\alpha(y)}$$

Clearly, one can check that

$$\lambda_n(\mathcal{A}_n) = F_*^n \nu(\mathcal{M}) = \mu(g(F^{-n})).$$

In particular, note that if we start from a standard pair (W, μ_W), then $F^n(W, \mu_W)$ is a standard family on $F^n W$ with measure $F^n \mu_W$ satisfying

$$F_*^n \mu_W(A) = \sum_{\alpha \in \mathcal{A}_n} \mu_\alpha(A \cap V_\alpha)\, \lambda_{\mathcal{M}}^u(\alpha).$$

Thus we use (3.23) and divide $\mu_\alpha(g \circ F^n)$ on every term to get

$$|g_\alpha(x) - g_\alpha(y)| \leq C_F |F^{-n} V_\alpha|^{\gamma_0} g_\alpha(y). \tag{3.25}$$

Since the probability density g_α is continuous on V_α, and the average value is 1, then there exists $y \in W_\alpha$, such that $g_\alpha(y) = 1$. Then we get that for any $x \in V_\alpha$:

$$|g_\alpha(x) - 1| \leq C_F |F^{-n} V_\alpha|^{\gamma_0},$$

where the constant C_F does not depend on x and y. Now for any Borel measurable set $A \subset \mathcal{M}$, we integrate the above inequality on A with respect to μ_α in order to get

$$|\int_{V_\alpha \cap A} g_\alpha \, d\mu_\alpha - \int_{V_\alpha \cap A} d\mu_\alpha| \leq C_F |F^{-n} V_\alpha|^{\gamma_0} \int_{V_\alpha \cap A} d\mu_\alpha. \tag{3.26}$$

Now we define a new standard family

$$\mathcal{G}_n^u := (F^n W, \mu_n) = ((V_\alpha, \mu_\alpha), \alpha \in \mathcal{A}_n, \lambda_n),$$

using the same factor measure, such that

$$\mu_n(A) = \sum_{\alpha \in \mathcal{A}_n} \lambda_n(\alpha) \int_{V_\alpha} \chi_A \, d\mu_\alpha.$$

We integrate (3.26) with respect to the factor measure λ_n, then it follows that

$$|F_*^n \nu(A) - \mu_n(A)| \leq C_F F_*^n \nu(A) \cdot \sup_{\alpha \in \mathcal{A}_n, V_\alpha \cap A \neq \emptyset} \{|F^{-n} V_\alpha|^{\gamma_0}\}.$$

This implies (3.21).

In general, we assume

$$\mathcal{G} = (\mathcal{W}, \nu) = \{(W_\alpha, \nu_\alpha), \alpha \in \mathcal{A}, \lambda\}$$

According to the above analysis, for any $n \geq 1$, and $\alpha \in \mathcal{A}$, $F^n(W_\alpha, \nu_\alpha)$ and its associated standard family, denoted as

$$\mathcal{G}_{n,\alpha}^u = (F^n W_\alpha, \mu_{n,\alpha}),$$

are related by

$$|F_*^n\nu_\alpha(A) - \mu_{n,\alpha}(A)| \le C_F\mu_{n,\alpha}(A)\cdot \sup_{\beta\in\mathcal{A}_{n,\alpha}, V_\beta\cap A\neq\emptyset}\{|F^{-n}V_\beta|^{\gamma_0}\}.$$

Now we integrate the above expression with respect to the factor measure, and using the fact that the associated family $\mathcal{G}_n^u$ of $F^n\mathcal{G}$ can be written as $\mathcal{G}_n^u = (F^n\mathcal{W}, \mu_n)$, which satisfies

$$\mu_n(A) = \int_{\alpha\in\mathcal{A}} \mu_{n,\alpha}(A)\lambda(d\alpha) = F_*^n\nu(A).$$

Thus for any Borel measurable set $A\subset\mathcal{M}$, we have

$$\begin{aligned}|F_*^n\nu(A) - \mu_n(A)| &= |\int_{\alpha\in\mathcal{A}} (F_*^n\nu_\alpha(A) - \mu_{n,\alpha}(A))\,\lambda(d\alpha)| \\ &\le C_F F_*^n\nu(A)\cdot \sup_{\beta\in\mathcal{A}_n, V_\beta\cap A\neq\emptyset}\{|F^{-n}V_\beta|^{\gamma_0}\},\end{aligned}$$

here $\mathcal{A}_n = \{W_\alpha \in F^{-n}A\}$ is the index set of unstable foliation of $F^{-n}A$, and $\mathcal{A}$ is the index set of $\mathcal{W}$. Now by (h1), i.e., the uniform expanding property of unstable manifolds under F leads to the claimed estimations. □

4 Coupling Lemma for the induced map

It was proved in [17, 32] that Assumptions **(h1)-(h4)** imply exponential decay of correlation for the induced system (M, F, μ_M) and any observables $f\in\mathcal{H}^-(\gamma_1)$ and $g\in\mathcal{H}^+(\gamma_2)$, where $\gamma_1,\gamma_2\in[\gamma_0,1]$, with $\operatorname{supp} f\subset M$ and $\operatorname{supp} g\subset M$. More precisely, we have

$$\left|\int_M (f\circ F^n)\,g\,d\mu_M - \int_M f\,d\mu_M\int_M g\,d\mu_M\right| \le C\|f\|^-_{C^{\gamma_1}}\|g\|^+_{C^{\gamma_2}}\vartheta^n, \tag{4.1}$$

for some uniform constants $\vartheta = \vartheta(\gamma_1,\gamma_2)\in(0,1)$ and $C>0$.

We will review in this section the coupling method developed in [39, 20, 26, 72] for the induced system, but we have to construct a special hyperbolic set.

4.1 Construction of a hyperbolic set $\mathcal{R}^*$

We first construct a hyperbolic set $\mathcal{R}^*\subset M$ with positive measure, which will be used as the reference set for the coupling procedure.

Definition 4.1. Let Γ^s be a family of stable manifolds, and Γ^u a family of unstable manifolds. We say that $\mathcal{R}^* = \Gamma^u\cap\Gamma^s$ is a *hyperbolic set with product structure*, if it satisfies the following four conditions:
(i) There exist a family of stable manifolds $\hat\Gamma^s$, a family of unstable manifolds $\hat\Gamma^u$, and a region $\mathcal{U}^*$ bounded by two stable manifolds $W_i^s\in\hat\Gamma^s$ and two unstable manifolds $W_i^u\in\hat\Gamma^u$, for $i=1,2$;

(ii) Any stable manifold $W^s \in \hat{\Gamma}^s$ and any unstable manifold $W^u \in \hat{\Gamma}^u$ only intersect at exactly one point;
(iii) The two defining families $\Gamma^{s/u}$ are obtained by intersecting $\hat{\Gamma}^{s/u}$ with $\mathcal{U}^*$, such that

$$\Gamma^{u/s} := \hat{\Gamma}^{u/s} \cap \mathcal{U}^*$$

(iv) Let $\nu^u = \mu|_{\Gamma^u}$ be obtained by restricting the SRB measure on Γ^u, then (Γ^u, ν^u) defines a standard family, and $\nu^u(\Gamma^s) > 0$.

We say a stable or unstable curve W *properly across* $\mathcal{U}^*$, if the two end points of the closure of $W \cap \mathcal{U}^*$ are contained in the boundary $\partial\mathcal{U}^*$. We say a set $A \subset \mathcal{R}^*$ is a u-subset, if there exists a measurable collection of unstable manifolds $\Gamma^u_A \subset \Gamma^u$, such that $A = \Gamma^u_A \cap \Gamma^s$. Similarly a set $A \subset \mathcal{R}^*$ is called a s-subset, if there exists a subset $\Gamma^s_A \subset \Gamma^s$, such that $A = \Gamma^s_A \cap \Gamma^u$.

It follows from condition (iv) that we can define a factor measure λ on the sigma algebra (induced by the Borel σ-algebra of $\mathcal{M}$) of the index set of $\Gamma^u = \{W_\alpha, \alpha \in \mathcal{A}\}$, such that for any Borel set $A \subset \mathcal{U}^*$,

$$\nu^u(A) = \int_{\alpha \in \mathcal{A}} \mu_\alpha(W_\alpha \cap A)\, \lambda(d\alpha).$$

Hyperbolic product sets were constructed in several references, see for example [39, 20, 26, 71, 32]. The next proposition was essentially proved in in [26, 32], so we will not repeat it here.

Proposition 4.2. There exist $\hat{\delta}_1$, a hyperbolic set with product structure $\mathcal{R}^* = \Gamma^s \cap \Gamma^u$ and the rectangle $\mathcal{U}^*$ containing $\mathcal{R}^*$ bounded by two stable manifolds and two unstable manifolds with length approximately $10\delta_0$, such that the following properties hold:
(i) $\mu(\mathcal{R}^*) > \hat{\delta}_1$ and for any unstable W that fully crosses $\mathcal{U}^*$, $\mu_W(\mathcal{R}^* \cap W) > \hat{\delta}_1$;
(ii) There exists $n_0 > 1$, such that for any $n \geq n_0$,

$$\mu(F^n \mathcal{R}^* \cap \mathcal{R}^*) > \hat{\delta}_1.$$

From now on, according to the construction in Proposition 4.2, we will fix the hyperbolic set $\mathcal{R}^*$, as well as its defining families Γ^s and Γ^u, with

$$\mathcal{R}^* = \Gamma^u \cap \Gamma^s. \tag{4.2}$$

In the coupling scheme that will be described below, we will consider a standard pair (W, ν) by subtracting from its density a smooth function. Next lemma explains that after a few more iterations under F, the resulting measure also induces a standard pair.

Lemma 4.3. *Let (W, ν) be a standard pair properly crosses $\mathcal{R}^*$, with $\rho = d\nu/d\mu_W$. Assume $g \in \mathcal{H}^+(\gamma_0)$ such that $\rho/3 < g < \rho/2$, we denote η as the measure with density $\rho_1 = \rho - g$. Then there exists a uniform constant $N \geq 1$, such that for any $n \geq N$, $F^n(W, \eta)$ is a proper family. Moreover, $(W, \eta/\eta(\mathcal{M}))$ is a pseudo-standard pair.*

Proof. By the definition of standard pair, we know that the positive density function $\rho = d\nu/d\mu_W$ satisfies (3.4):

$$|\ln\rho(x) - \ln\rho(y)| \le C_F d(x,y)^{\gamma_0}, \tag{4.3}$$

where $\gamma_0 \in (0,1)$ was given in (2.5), and $C_F > C_{\mathbf{r}}$ is a fixed large constant. Now for $g \in \mathcal{H}^+(\gamma_0)$, with $\rho/3 < g < \rho/2$, we denote

$$\rho_1 = \rho - g$$

as the density of η. According to (3.9), one can check that for any measurable set A,

$$F_*^n\eta(A) = \int_{\alpha\in\mathcal{A}_n}\int_{V_\alpha}\chi_A(x)d\nu_\alpha\,\lambda_n(d\alpha),$$

where

$$d\nu_\alpha = \rho_\alpha\,d\mu_\alpha$$

and

$$\lambda_n(d\alpha) = \left(\int_{V_\alpha}\rho_1(F^{-n}x)\,d\mu_\alpha(x)\right)\lambda_M^u(d\alpha)$$

is the factor measure on index set $\mathcal{A}_n$. Here ρ_α is the probability density function defined only on V_α such that for any $x \in V_\alpha$,

$$\rho_\alpha(x) = \frac{\rho_1(F^{-n}x)}{\int_{V_\alpha}\rho_1(F^{-n}y)d\mu_\alpha(y)}.$$

Note that using the fact that $\rho/3 \le g \le \rho/2$, we have for any $x, y \in V_\alpha \in F^nW$,

$$\begin{aligned}|\ln\rho_\alpha(x) - \ln\rho_\alpha(y)| &= |\ln\frac{\rho(F^{-n}x) - g(F^{-n}x)}{\rho(F^{-n}y) - g(F^{-n}y)}| \\ &\le 2C_F d(F^{-n}x, F^{-n}y)^{\gamma_0} \le 2C_{\mathbf{r}}C_F\Lambda^{-n\gamma_0}d(x,y)^{\gamma_0},\end{aligned}$$

where $C_{\mathbf{r}} > 0$ is the distortion constant, and $\Lambda > 1$ is the minimal expansion factor for F. Thus we choose the smallest $N > 1$, such that

$$2C_{\mathbf{r}}\Lambda^{-N\gamma_0} < 1$$

Then combining with Lemma 3.5, for any $n \ge N$, $F^n(W,\eta)$ is a proper family. If we define $\eta' = \eta/\eta(\mathcal{M})$, then one can check that (W,η') is indeed a pseudo-standard pair. □

Definition 4.4. Let $\mathcal{K} = \{(W_\alpha,\nu_\alpha), \alpha \in \mathcal{A}\}$ be a family of pseudo-standard pairs that fully cross $\mathcal{R}^*$, equipped with a factor measure λ. Then we call $\mathcal{K}$ a *pseudo-proper family.*

Lemma 4.5. *There exist $\hat{\delta}_0 \in (0,\mu(\Gamma^s))$, $N_0 > 1$, such that for any proper family $\mathcal{G} = (\mathcal{W},\nu)$ with $\nu(M) = 1$, then $F_*^n\nu$ has at least $\hat{\delta}_0$ portion of measure properly returned to $\mathcal{R}^*$, for any $n \ge N_0$.*

Proof. Let $\hat{\delta}_1 > 0$ be defined as in Proposition 4.2, such that $\mu(\mathcal{R}^*) > \hat{\delta}_1$. By the uniform mixing property of the induced map (F, M, μ_M), and the fact that $\mathcal{G}$ is a proper family, (4.1) implies that for $n > 1$,

$$|F_*^n \nu^i(\mathcal{R}^*) - \mu_M(\mathcal{R}^*)| \leq C\vartheta^n.$$

Moreover for any standard pair (W_α, ν_α), since W_α only has two end points, say x_1, x_2, so if $F^n W_\alpha$ intersects $\mathcal{R}^*$ at some $x \in \mathcal{R}^*$, then it must consist of the entire unstable manifold $W^u(x)$, unless $W^u(x)$ consists of one of points in $\{F^n x_1, F^n x_2\}$. Thus a majority of curves in $F^n\mathcal{W}$ must properly cross Γ^s.

Thus by taking a large N_0 and a small number $\hat{\delta}_0 \in (0, \hat{\delta}_1)$, we have that for any $n \geq N_0$, $F_*^n \nu$ has at least $\hat{\delta}_0$ portion of measure properly returned to $\mathcal{R}^*$. Moreover, our choice of N_0 and $\hat{\delta}_0$ are uniform for all proper families.

□

Let $N \geq 1$ be the integer chosen in Lemma 4.3. We define $n_1 = \max\{N, N_0\}$, and we define a higher iteration

$$\tilde{F} := F^{n_1} \tag{4.4}$$

From now on we consider $\tilde{F}$ instead of F. Note that by Lemma 4.3, if we subtract a "nice" function from the density of a proper standard pair (W, ν), then after one iteration of $\tilde{F}$, the image $\tilde{F}(W, \eta)$ becomes a new proper family, where η is the new conditional measure and has at least $\hat{\delta}_0$ portion of measure properly returned to $\mathcal{R}^*$.

4.2 The Coupling Lemma for the induced map

Next we restate the Coupling Lemma [20, 26] for the induced system $(\tilde{F}, \mu_M)$ using the concept of standard families.

Lemma 4.6. *Under assumption **(h1)-(h4)**. Let $\mathcal{G}^i = (\mathcal{W}^i, \nu^i), i = 1, 2$, be two standard families on M. For any $n \geq 1$, there exist two sequences of standard families $\{(\mathcal{W}_k^i, \nu_k^i), k = 1, \cdots, n\}$, and two standard families $\{(\bar{\mathcal{W}}_n^i, \bar{\nu}_n^i)\}$, such that for any measurable collection of stable manifolds $A \subset \mathcal{W}^s$,*

$$F_*^n \nu^i(A) = \sum_{k=1}^{n} F_*^{n-k} \nu_k^i(A) + \bar{\nu}_n^i(A).$$

And they also have the following properties, for each $k = 1, \cdots, n$:

(i) ***Proper returned to $\mathcal{R}^*$ at k.***
Both $\mathcal{W}_k^1$ and $\mathcal{W}_k^2$ are u-subsets of Γ^u;

(ii) ***Coupling ν_k^1 and ν_k^2 along stable manifolds in Γ^s.***
For any measurable collection of stable manifolds $A \subset \Gamma^s$, we have

$$\nu_k^1(A) = \nu_k^2(A)$$

(iii) ***Exponential tail bound for uncoupled measure at*** n***.***
There exists $N = N(\nu^1, \nu^2) \geq 1$*, such that for any* $n > N$*,*

$$\bar{\nu}_n^i(M) < C\vartheta^n, \tag{4.5}$$

where $C > 0$ *and* ϑ *are uniform constants. If both* $\mathcal{G}^1$ *and* $\mathcal{G}^2$ *are proper, then we can take* $N = 1$*.*

Note that the original Coupling Lemma was stated only for proper families, so to deal with a standard family which is not proper, we need to iterate N times to make it proper, according to Lemma 3.5 and the remark before it. Here $\nu_n^i, i = 1, 2$, are the coupled components of $\tilde{F}_*^n \nu^i$. In practice, a coupling procedure occurs at a sequence of times $0 < t_1 < t_2 < \cdots < t_k < \infty$. In particular, $\nu_j^i = 0$, when $j \neq t_k$ for all $1 \leq k$, which means that ν_j^i remains unchanged between successive coupling times.

According to the above Lemma, for any bounded function f that is constant on each stable manifold, we have

$$\begin{aligned} |\tilde{F}_*^n \nu^1(f) - \tilde{F}_*^n \nu^2(f)| &\leq \sum_{k=1}^{n} |\tilde{F}_*^{n-k}(\nu_n^1(f) - \nu_n^2(f))| + |\bar{\nu}_n^1(f) - \bar{\nu}_n^2(f)| \\ &= |\bar{\nu}_n^1(f) - \bar{\nu}_n^2(f)| \leq 2C\|f\|_\infty \vartheta^n. \end{aligned} \tag{4.6}$$

Similarly, the above coupling lemma implies the exponential rates for any bounded Hölder function $f \in \mathcal{H}^-(\gamma_1)$ for the induced system $(\tilde{F}, M)$:

$$\begin{aligned} &\left| \int f \circ \tilde{F}^n \, d\nu^1 - \int f \circ \tilde{F}^n \, d\nu^2 \right| \\ &\leq 2\|f\|_\infty \bar{\nu}_n^1(M) + \sum_{j \leq n} \|f\|_{\gamma_1}^- \Lambda^{-(n-j)\gamma_1} \nu_j^1(M) \\ &\leq 2C\|f\|_{C^{\gamma_1}}^- \vartheta_2^n, \end{aligned} \tag{4.7}$$

where $\vartheta_2 = \min\{\Lambda^{-\gamma_1}, \vartheta\}$.

Next we prove a lemma which directly follows from the above Coupling Lemma 4.6.

Lemma 4.7. *For any standard family* $\mathcal{G} = (\mathcal{W}, \nu)$ *in* M*, there exists a sequence of standard families* $\{(\mathcal{V}_n, \eta_n), n \geq 1\}$ *with the following properties:*

(i) $\tilde{F}^n(\mathcal{V}_n)$ *is a u-subset of* Γ^u*;*

(ii) For any measurable collection of stable manifolds $A \subset \mathcal{W}^s$*,*

$$\nu(A) = \sum_{n \geq 0} \eta_n(A \cap \mathcal{W}_n^s)$$

where

$$\mathcal{W}_n^s = \{W^s(x) \,:\, \tilde{F}^n x \in \mathcal{W}_n \cap \Gamma^s\} \subset \mathcal{V}_n \cap \tilde{F}^{-n}\Gamma^s;$$

(iii) Furthermore there exists $N = N(\nu) \geq 1$, such that for any $n > N$,

$$\sum_{k=n}^{\infty} \tilde{F}_*^n \nu_k(\mathcal{W}_k^s) < C\vartheta^n, \tag{4.8}$$

where $C > 0$ and ϑ is the constant in (4.5).

Proof. For $i = 1, 2$, we take

$$\mathcal{G}^i := (\mathcal{W}, \nu)$$

as two copies of the standard family. Then according to the Coupling Lemma 4.6, by taking $n \to \infty$, there exists a sequence of standard families $\{(\mathcal{W}_n, \nu_n), n \geq 1\}$, such that (1)-(3) hold (as stated in Lemma 4.6). Moreover, one can check that the uncoupled measure at time n satisfies:

$$\bar{\nu}_n(M) = \sum_{k \geq n} \tilde{F}_*^n \nu_k(\Gamma^s).$$

We define $(\mathcal{V}_n, \eta_n)$, such that

$$\tilde{F}^n \mathcal{V}_n = \mathcal{W}_n \quad \text{and} \quad \tilde{F}_*^n \eta_n = \nu_n.$$

Then (1)-(3) imply our statement (i)-(ii). In particular, we have for any measurable set $A \subset M$,

$$\nu(A) = \sum_{n=1}^{\infty} \eta_n(A),$$

which implies our third statement:

$$\sum_{k=n}^{\infty} \tilde{F}_*^n \nu_k(\Gamma^s) = \bar{\nu}_n(M) \leq C\vartheta^n.$$

□

Next we will show that there is a generalized Young tower on M, following the above Coupling Lemma. The proof of the existence of a generalized Young tower as a consequence of the Coupling Lemma was first derived in [32] implicitly, and also proved in [70].

Proposition 4.8. By choosing n_1 large enough in (4.4), the induced map $\tilde{F}$ defines a generalized countable Markov partition of $\mathcal{R}^*$ into s-subsets

$$\mathcal{R}^s = \cup_{n \geq 1} \hat{\mathcal{R}}_n^s$$

with the following properties:

(a) For any $n \geq 1$, if $\tilde{F}^n(\hat{\mathcal{R}}_n^s)$ is nontrivial, then it returns to $\mathcal{R}^*$, and $\tilde{F}^n(\hat{\mathcal{R}}_n^s)$ is a u-subset of $\mathcal{U}^*$. Moreover, there exists $\delta' > 0$ such that

$$\mu(\hat{\mathcal{R}}_1^s) > \delta'$$

and

$$\sum_{k=n}^{\infty} \mu(\hat{\mathcal{R}}_k^s) < C\vartheta^n. \tag{4.9}$$

(b) For each $n \geq 1$, there exists at most countably many s-subset $\hat{\mathcal{R}}^s_{n,i}$, $i \geq 1$, such that

$$\hat{\mathcal{R}}^s_n = \cup_{i\geq 1}\hat{\mathcal{R}}^s_{n,i}$$

and for any $W^u \in \Gamma^u$, any $x, y \in \hat{\mathcal{R}}^s_{n,i} \cap W^u$, we have $\tilde{F}^n W^u(x) \in \Gamma^u$ properly cross $\mathcal{U}^*$ and $\tilde{F}^n y \in \tilde{F}^n W^u(x)$.

Proof. We first construct a partition of $\mathcal{R}^*$ into s-subsets. In the Coupling Lemma 4.6, for $i = 1, 2$, we take

$$\mathcal{E}^i := (\Gamma^u, \mu|_{\Gamma^u}) = \{(W, \mu_W), W \in \Gamma^u, \lambda^u\}$$

as two copies of the proper family induced by the SRB measure μ restricted on the family of all unstable manifolds Γ^u, with factor measure λ^u. Then at each iteration $n \geq 1$, we couple everything in $\tilde{F}^n\mathcal{E}^i$ that properly returned to $\mathcal{R}^*$. More precisely, Lemma 4.7 implies that there exists a sequence of standard families $\{(\mathcal{V}_k, \eta_k), k \geq 1\}$, such that

$$\tilde{F}^k(\mathcal{V}_k, \eta_k) = (\mathcal{W}_k, \nu_k),$$

where $\mathcal{W}_k$ is a u-subset of Γ^u. Moreover, we have

$$\nu(\Gamma^s) = \sum_{k\geq 1} \eta_k(\Gamma^s \cap \mathcal{W}^s_k),$$

where $\mathcal{W}^s_k = \{W^s(x) : \tilde{F}^k x \in \Gamma^s \cap \mathcal{W}_k\}$ is a collection of stable manifolds that intersect $\mathcal{V}_k$. We define

$$\hat{\Gamma}^s_k = \mathcal{W}^s_k \cap \Gamma^s.$$

Then we can check that

$$\Gamma^s = \cup_{n=1}^{\infty}\hat{\Gamma}^s_n \pmod 0$$

with the following properties:

(1) $\tilde{F}^n(\hat{\Gamma}^s_n)$ is a u-subset of $\mathcal{U}^*$ and $\{\hat{\Gamma}^s_n, n \geq 1\}$ are almost surely disjoint s-subsets of $\mathcal{U}^*$ in the following sense:

$$\mu(\hat{\Gamma}^s_m \cap \hat{\Gamma}^s_n) = 0$$

for any $m \neq n$;

(2) Furthermore

$$\sum_{k=n}^{\infty} \mu(\hat{\Gamma}^s_k) < C\vartheta^n, \tag{4.10}$$

where $C > 0$ and ϑ is the constant in (4.5).

The fact that a nonempty set $\tilde{F}^n(\hat{\Gamma}^s_n)$ properly return to $\mathcal{R}^*$ is guaranteed by Proposition 4.2. By taking n_1 large enough, Lemma 4.5 implies the existence of $\delta' \in (0, \hat{\delta}_0)$ such that

$$\mu(\hat{\Gamma}^s_1) > \delta'.$$

We next verify (b). Because our singular set $S_{\pm 1}$ contains at most countably many smooth curves, it is possible that for an unstable manifold $W \in \Gamma^u$, its image $\tilde{F}^n W$ contains countably many smooth components that properly returned to $\mathcal{U}^*$. Thus there exist countably many s-subsets $\hat{\Gamma}^s_{n,i}$, $i \geq 1$, such that

$$\hat{\Gamma}^s_n = \cup_{i \geq 1} \hat{\Gamma}^s_{n,i},$$

and for any $x \in \hat{\Gamma}^s_{n,i}$, $\tilde{F}^n W^u(x) \in \Gamma^u$ properly crosses $\mathcal{U}^*$. Moreover, for any $x, y \in \hat{\Gamma}^s_{n,i} \cap W^u$, $\tilde{F}^n y$ and $\tilde{F}^n x$ belong to the same unstable manifold in Γ^u.

We define

$$\hat{\mathcal{R}}_k = \mathcal{R}^* \cap \hat{\Gamma}^s_k$$

and thus

$$\mathcal{R}^* = \cup_{n \geq 1} \hat{\mathcal{R}}_k$$

with property (a) and (b). □

Note that one can easily build up the generalized Young Tower based on the Markov partition $\mathcal{R}^* = \cup_{n \geq 1} \hat{\mathcal{R}}_n$ in the spirit of [71, 72]. In addition, we have a partition of the phase space M:

$$M = \bigcup_{n \geq 1} \bigcup_{k=0}^{n-1} F^k \hat{\mathcal{R}}_n \ \ (\bmod\ 0) \tag{4.11}$$

One improvement here is that according to statement (b) in the above lemma, we allow the minimal s-rectangle containing $\hat{\Gamma}^s_n$ to consist of countably many minimal s-rectangles $\mathcal{U}^*_{n,i}$. This property is due to the fact that we allow the singular set of the system to contain countably many singular curves, since one unstable manifold may be cut into infinitely many small pieces, many of which may returned to the rectangle $\mathcal{U}^*$ simultaneously. To model general systems with countable singularities, a generalized Young tower with property (b) is indeed required.

Next we investigate the relation between the set $C_{n,b}$ defined in (2.2) and the reference set $\mathcal{R}^*$. Indeed by Assumption (**H2**), we know that

$$\mu(M \cap C_{n,b}) = \mathcal{O}(n^{1-\alpha_0-\varepsilon_1}).$$

We would like to see similar property for standard pairs that properly cross $\mathcal{R}^*$.

Lemma 4.9. *There exists $c_0 > 0$, such that for any standard pair (W, ν) properly cross $\mathcal{R}^*$, with $d\nu/d\mu_W = g \in \mathcal{H}^+(\gamma_0)$, then for any $n \geq 1$,*

$$\nu(C_{n,b} \cap W \cap \mathcal{R}^*) \leq c_0 \|g\|^+_{C^{\gamma_0}} n^{1-\alpha_0-\varepsilon_1}. \tag{4.12}$$

Proof. Since Lemma 4.2 implies that $\mu(\Gamma^s) > \hat{\delta}_1$, by Assumption (**H2**), for any $n \geq 1$,

$$\mu(\Gamma^s \cap C_{n,b}) \leq \mu(C_{n,b} \cap M) \leq Cn^{1-\alpha_0-\varepsilon_1},$$

for some constant $C > 0$. One can check that for $x \in C_{n,b}$, then its stable manifold $W^s(x) \in C_{n,b}$, and

$$\mu(W^s \in C_{n,b} \cap \Gamma^s) = \mu(\Gamma^s \cap C_{n,b}) \leq Cn^{1-\alpha_0-\varepsilon_1}.$$

Now we disintegrate the measure μ restricted on $\mathcal{U}^*$ along unstable leaves in $\Gamma^u = \{W_\alpha, \alpha \in \mathcal{A}\}$, and let λ be the factor measure on the index set $\mathcal{A}$, such that $\lambda(\mathcal{A}) = \mu(\Gamma^s)$, and for any measurable set A,

$$\mu(\Gamma^s \cap A) = \int_\alpha \mu_\alpha(W_\alpha \cap A \cap \Gamma^s)\, \lambda(d\alpha).$$

Picking a curve $W_{\alpha_1} \in \Gamma^u$, by the absolute continuity of the stable holonomy map **(h3)**, there exist $0 < c_1 < c_2$, such that for any $\alpha \in \mathcal{A}$, any measurable set $A \subset \mathcal{M}$ satisfies

$$\begin{aligned} c_1\mu_{\alpha_1}(W_{\alpha_1} \cap A \cap \Gamma^s) &\leq \mu_\alpha(W_\alpha \cap A \cap \Gamma^s) \\ &\leq c_2\mu_{\alpha_1}(W_{\alpha_1} \cap A \cap \Gamma^s) \end{aligned}$$

This implies that

$$\mu_{\alpha_0}(W_{\alpha_1} \cap A \cap \Gamma^s)\lambda(\mathcal{A}) \leq c_1^{-1}\mu(\Gamma^s \cap A).$$

Now we take any unstable manifold $W \in \Gamma^u$, and $A = C_{n,b}$, then we have proved that

$$\mu_W(W \cap C_{n,b} \cap \Gamma^s) \leq Cc_1^{-1}\hat{\delta}_1^{-1}n^{1-\alpha_0-\varepsilon_1}.$$

Since unstable manifolds in Γ^s has length $> 10\delta_0$, Lemma 3.5 implies that a standard pair (W, ν) is proper whenever W cross Γ^s. So (W, ν) and (W, μ_W) are equivalent:

$$\nu(W \cap C_{n,b} \cap \Gamma^s) \leq C_1Cc_1^{-1}\delta_1^{-1}\|g\|^+_{C^{\gamma_0}}n^{1-\alpha_0-\varepsilon_1},$$

for some constant C_1 depending on C_F in (3.4). Now we take

$$c_0 = C_1Cc_1^{-1}\delta_1^{-1},$$

then (4.12) has been proved. □

5 Markov tower for the original map

5.1 Construction of the Markov tower for the original map

In this subsection, we will construct a countable Markov partition of Γ^s for the nonuniformly hyperbolic map and then use the return time to $\mathcal{U}^*$ to define a

stopping time for our coupling scheme. To investigate the map $(\mathcal{F}, \mathcal{M}, \mu)$ based on the induced system $(\tilde{F}, M, \mu_M)$, we know that $\tilde{F}$ and $\mathcal{F}$ share the same stable/unstable manifolds on M almost surely. This allows us to use the same reference set $\mathcal{R}^*$ and $\mathcal{U}^*$, as well as the stable/unstable manifolds $\Gamma^{s/u}$ that defines $\mathcal{R}^*$. First we extend the partition according to the original map by the following construction.

Proposition 5.1. There exists $n_2 \geq n_1$, and we denote $\tilde{\mathcal{F}} := \mathcal{F}^{n_2}$, with the following properties:

(i) For any $n \geq 1$, the set
$$\hat{\mathcal{R}}_n = \bigcup_{m \geq n} \mathcal{R}_{n,m}$$
has a decomposition into s-subsets $\mathcal{R}_{n,m}$ such that for any fixed $m \geq n$, $\tilde{F}^n \mathcal{R}_{n,m}$ returns to $\mathcal{R}^*$ and is a u-subset of $\mathcal{U}^*$; moreover $\tilde{F}^n \mathcal{R}_{n,m} = \tilde{\mathcal{F}}^m \mathcal{R}_{n,m}$;

(ii) $\mathcal{R}^*$ has a partition into s-subsets $\mathcal{R}^* = \cup_{n \geq 1} \mathcal{R}_n$, such that if $\tilde{\mathcal{F}}^n \mathcal{R}_n$ is nonempty, then it is a u-subset of $\mathcal{U}^*$;

(iii) There exist $\delta_2 \in (0, \hat{\delta}_0)$ and $C > 0$, such that for any proper family $(\mathcal{W}, \nu)$ with $\mathcal{W} \subset \Gamma^u$ and $\nu(\mathcal{U}^*) = 1$, then $\nu(\mathcal{R}_1) > \delta$ and
$$\nu(\cup_{m \geq n} \mathcal{R}_m) \leq C n^{-\alpha_0}$$
for any $n > 1$.

Proof. By Proposition 4.8, the set Γ^s has a countable Markov partition into s-subsets
$$\Gamma^s = \cup_{n \geq 1} \hat{\Gamma}^s_n.$$

Statement (a) of Proposition 4.8 implies that $\mu(\hat{\Gamma}^s_1) > 0$. Let
$$A_1 = \tilde{F} \hat{\Gamma}^s_1,$$
then A_1 is a u-subset. Note that by definition of $\tilde{F} = F^{n_1}$, we know that for almost any $x \in \hat{\Gamma}^s_1$, there exists $m = m(x) \geq n_1$, such that $\tilde{F}^n x = \mathcal{F}^m x$. Let $W(x)$ be the unstable manifold of x, then $\mathcal{F}^m W(x)$ properly cross Γ^s. Moreover there exists a minimal u-subset in A_1, denoted as $A_{1,m}$, containing $\mathcal{F}^m W(x) \cap \Gamma^s$, such that $\mathcal{F}^{-m} A_{1,m}$ is a s-subset of $\mathcal{U}^*$. This implies that there exists a countable decomposition of
$$\hat{\Gamma}^s_1 = \bigcup_{m \geq n_1} \mathcal{F}^{-m} A_{1,m}$$
into s-subsets. Clearly, there exist $n_2 \geq n_1$ and $\delta_0' \in (0, \hat{\delta}_0)$, such that
$$\Gamma_{1,1} := \mathcal{F}^{-n_2} A_{1,n_2}$$
and
$$\mu(\Gamma_{1,1}) > \delta_0'.$$

Thus we define $\tilde{\mathcal{F}} = \mathcal{F}^{n_2}$, and we define $\Gamma_{1,m}$ according to the number of iterations under $\tilde{\mathcal{F}}$, such that $\tilde{\mathcal{F}}^m \Gamma_{1,m}$ is a u-subset, and

$$\hat{\Gamma}_1^s = \cup_{n=1}^{\infty} \Gamma_{1,m}.$$

Inductively for any $n \geq 1$, and $m \geq n$, we can define $\Gamma_{n,m}$, which is the maximal s-subset of Γ^s, such that $\tilde{\mathcal{F}}^m \Gamma_{n,m}$ is a u-subset of $\mathcal{U}^*$. Moreover $\Gamma_{n,m}$ will also properly return to $\mathcal{U}^*$ under the induced map $\tilde{F}^n$. Then it follows that

$$\hat{\Gamma}_n^s = \bigcup_{m \geq n} \Gamma_{n,m}$$

is a disjoint decomposition of $\hat{\Gamma}_n^s$, for any $n \geq 1$. Next we rearrange $\{\Gamma_{n,m}\}$ according to the index m. Note that

$$\Gamma^s = \bigcup_{n \geq 1} \hat{\Gamma}_n^s = \bigcup_{n \geq 1} \bigcup_{m \geq n} \Gamma_{n,m} = \bigcup_{m=1}^{\infty} \left(\bigcup_{n=1}^{m} \Gamma_{n,m} \right) = \bigcup_{m=1}^{\infty} \Gamma_m^s$$

where

$$\Gamma_m^s = \cup_{n=1}^{m} \Gamma_{n,m}.$$

Then by the definition of $\Gamma_{n,m}$, we know that $\tilde{\mathcal{F}}^m \Gamma_m^s$ returns properly to $\mathcal{R}^*$, and is nonempty. Moreover,

$$\mu(\Gamma_{1,1}) > \delta_0'$$

implies that for any proper family $(\mathcal{W}, \nu)$ with $\mathcal{W} \subset \Gamma^u$ and $\nu(\mathcal{U}^*) = 1$, then

$$\nu(\Gamma_1^s) > \delta_2$$

for some $\delta_2 \in (0, \delta_0')$.

Next we estimate the measure of Γ_m^s.

By our definition of $\hat{\Gamma}_1^s$ and the choice of n_1 in (4.4), we know that $\tilde{F}\Gamma_{1,m} = \tilde{\mathcal{F}}^m \Gamma_{1,m}$. Thus

$$\begin{aligned}
\mu_M(\cup_{m=n}^{\infty} \Gamma_m^s) &\leq \mu_M(R > n) + \mu_M(\cup_{m=n}^{\infty} \Gamma_m^s, R < n, C_{n,b}) + \mu_M(\cup_{m=n}^{\infty} \Gamma_m^s, R < n, C_{n,b}^c) \\
&\leq \mu_M(R > n) + \mu_M(\cup_{m=n}^{\infty} \Gamma_m^s, R < n, C_{n,b}^c) + \mu_M(R < n, C_{n,b}) \\
&= \mu_M(R > n) + \mathcal{O}(n^{1-\alpha_0-\varepsilon_1}) + \mathcal{O}(n^{-1-\alpha_0}) \\
&= \mu_M(R > n) + \mathcal{O}(n^{1-\alpha_0-\varepsilon_1}).
\end{aligned}$$

Here in the last step of the above estimations, we have used the fact that the set

$$\{x \in M \,:\, \cup_{m=n}^{\infty} \Gamma_m^s, R < n, C_{n,b}^c\}$$

contains points that have returned to M at least $b \ln n$ times within n iterations, but have not been coupled (as they have not returned to $\mathcal{R}^*$ yet). According

to (4.8), we know that the standard family $(\mathcal{W}^u, \mu)$ has a decomposition into $\{(\mathcal{W}_n, \mu_n), n \geq 1\}$, and $\sum_{m=n}^{\infty} \mu_m(M) \leq C\vartheta^n$. Thus

$$\begin{aligned}\mu(\cup_{m=n}^{\infty}\Gamma_m^s, R < n, C_{n,b}^c) &= \sum_{m=b\ln n}^{\infty} \mu_m(R < n, C_{n,b}^c) \\ &\leq C\vartheta^{b\ln n} = \mathcal{O}(n^{-1-\alpha_0}).\end{aligned}$$

By the absolutely continuity property for the stable holonomy map, for any proper family $(\mathcal{W}, \nu)$ with $\mathcal{W} \subset \Gamma^u$ with $\nu(\mathcal{U}^*) = 1$, we have

$$\nu(\cup_{m\geq k}\Gamma_m^s) \leq Ck^{-\alpha_0},$$

for some $C > 0$.

Finally, we define $\mathcal{R}_m = \Gamma_m^s \cap \mathcal{R}^*$, and $\mathcal{R}_{m,n} = \Gamma_{m,n}^s \cap \mathcal{R}^*$. Therefore, the above analysis verifies items (i)-(iii). □

5.2 Proper returns to the base of the tower

In the next sections, we will also consider the images of a standard family $(\mathcal{W}, \nu)$ under iterations of the map $\mathcal{F}$. Indeed one can also show that $\mathcal{F}^n(\mathcal{W}, \nu)$ essentially becomes a proper family as long as n is large.

Lemma 5.2. *Fix any $\delta > 0$. Let $\mathcal{G} = (\mathcal{W}, \nu)$ be any standard family. Then there exists $N_1 \geq 1$, such that $\mathcal{F}^{N_1}\mathcal{G}$ is a proper family. Moreover, by picking N_1 large enough, $\mathcal{F}^{N_1}\mathcal{G}$ has at least δ portion of the measure that fully returned to $\mathcal{R}^*$.*

Proof. Since by the mixing property, we know that $\mathcal{F}_*^n\nu \to \mu$ weakly. Moreover, as $(\mathcal{W}^u, \mu)$ is a proper family, one can check that there exists $N_1 \geq 1$, such that $\mathcal{F}^n\mathcal{G}$ is also proper, as $n \geq 1$. The second statement also follows from the mixing property and the fact that $\mu(\mathcal{R}^*) > 0$. □

Remark 5.3. Since our goal in this paper is to investigate the polynomial mixing rates for $(\mathcal{F}, \mathcal{M}, \mu)$, the order of convergence will not alter even if we consider a higher iteration $\mathcal{F}^{n_2}$, as $(nn_2)^{-a} = n_2^{-a}n^{-a}$, which simply adds an extra factor in the coefficient of the decay rates. **From now on, we will only consider higher iterations $\tilde{\mathcal{F}} := \mathcal{F}^{n_2}$, $\tilde{F} := F^{n_1}$, but still denote them as $\mathcal{F}$ and F, respectively, just for simplicity of notations.** More precisely, this implies that the (new) map $\tilde{\mathcal{F}}$ satisfies Proposition 5.1, in particularly, we have $\mu(\Gamma_1^s) > \delta$; and $\tilde{F}$ satisfies Proposition 4.8, we have

$$\mu_M(\hat{\Gamma}_1^s) > \delta'.$$

Lemma 5.4. *Let $\mathcal{G} = (\mathcal{W}, \nu)$ be a standard family. Then there exists a sequence of standard families $\{(\hat{\mathcal{W}}_n, \hat{\nu}_n), n \geq 0\}$, such that if $\mathcal{F}^n(\hat{\mathcal{W}}_n) \cap \mathcal{R}^*$ is nontrivial, then it returns to $\mathcal{R}^*$, and is a u-subset of $\mathcal{U}^*$; moreover, for any measurable collection of stable manifolds $A \subset \mathcal{W}^s$, we have*

$$\nu(A) = \sum_{n\geq 0} \hat{\nu}_n(A \cap \mathcal{W}_n^s)$$

where

$$\mathcal{W}_n^s = \{W^s(x) \,:\, \mathcal{F}^n x \in \Gamma^s \cap \mathcal{F}^n \hat{\mathcal{W}}_n\}.$$

Proof. Let $R : \mathcal{M} \to \mathbb{N}$ be the first hitting time function defined as in (3.1), we define

$$\mathcal{W}^n := \mathcal{W} \cap ((R = n) \setminus M),$$

for $n \geq 1$, and

$$\mathcal{W}^0 = \mathcal{W} \cap M.$$

Note that the set $\mathcal{F}^n \mathcal{W}^n \subset M$. Now

$$\mathcal{E}_n := \mathcal{F}^n(\mathcal{W}^n, \nu|_{\mathcal{W}^n})$$

becomes a standard family on M. By Lemma 4.7, there exists a sequence of standard families $\{(V_{n,k}, \eta_{n,k}), k \geq 1\}$ associated with $\mathcal{E}_n$. More precisely, $F^k V_{n,k}$ is a u-subset of $\mathcal{U}^*$ and for any measurable collection A of stable manifolds,

$$\mathcal{F}_*^n \nu(A) = \sum_{k=1}^{\infty} \eta_{n,k}(A \cap \mathcal{W}_{n,k}^s),$$

where

$$\mathcal{W}_{n,k}^s = \{W^s(x) \,:\, F^k x \in \Gamma^s \cap F^k \mathcal{V}_{n,k}\}.$$

Moreover, we decompose $\{(V_{n,k}, \eta_{n,k}), k \geq 1\}$, according to its iterations under the original map $\mathcal{F}$, such that

$$F^k(V_{n,k,i}, \nu_{n,k,i}) = \mathcal{F}^i(V_{n,k,i}, \nu_{n,k,i}).$$

Here $V_{n,k,i}$ is the support of $\nu_{n,k,i}$, and $F^k V_{n,k,i} = \mathcal{F}^i V_{n,k,i}$ is a u-subset of $\mathcal{U}^*$. Let

$$\nu_{n,i} = \sum_{k=0}^{i} \nu_{n,k,i}$$

and

$$\mathcal{W}_{n,i} = \cup_{k=0}^{i} V_{n,k,i}.$$

Now we define

$$\hat{\nu}_k = \sum_{n=0}^{k} \mathcal{F}_*^{-n} \nu_{n,k-n}.$$

Let

$$\hat{\mathcal{W}}_k = \mathrm{supp}(\hat{\nu}_k)$$

then by the definition of $\hat{\nu}_k$, we know that $\mathcal{F}^k(\hat{\mathcal{W}}_k)$ is also an (union of) u-subset that properly returned to $\mathcal{U}^*$. This verifies the two statements as claimed. □

According to the above Lemma, we will introduce a new concept called the *generalized standard family.*

Definition 5.5. For any standard family $\mathcal{G} = (\mathcal{W}, \nu)$, let $\{(\hat{\mathcal{W}}_n, \hat{\nu}_n), n \geq 0\}$ be a sequence of standard families constructed in Lemma 5.4, such that $\mathcal{F}^n(\hat{\mathcal{W}}_n) \cap \mathcal{R}^*$ returns to $\mathcal{R}^*$, and is a u-subset of $\mathcal{U}^*$; and for any measurable collection of stable manifolds $A \subset \mathcal{W}^s$,

$$\nu(A) = \sum_{n \geq 0} \hat{\nu}_n(A \cap \mathcal{W}^s_n),$$

where

$$\mathcal{W}^s_n = \{W^s(x) \,:\, \mathcal{F}^n x \in \Gamma^s \cap \mathcal{F}^n \hat{\mathcal{W}}_n\}.$$

We call

$$(\mathcal{W}_n, \nu_n) := (\hat{\mathcal{W}}_n \cap \mathcal{W}^s_n, \hat{\nu}_n|_{\mathcal{W}^s_n})$$

a **generalized standard family with index** n and $(\hat{\mathcal{W}}_n, \hat{\nu}_n)$ its *shadow standard family.* We also denote

$$\mathcal{G} = \sum_{n \geq 0} (\mathcal{W}_n, \nu_n)$$

as *a standard decomposition* of $\mathcal{G}$. Moreover, if the shadow $(\hat{W}, \hat{\nu})$ is a pseudo-proper family with $\hat{W} \subset \Gamma^u$, then we call $(\hat{W}, \hat{\nu})|_{\mathcal{R}^*}$ a pseudo generalized standard family with index 0.

The advantage of making a such definition is that if we know $(\mathcal{W}_n, \nu_n)$ is a generalized standard family of index n with shadow $(\hat{\mathcal{W}}_n, \hat{\nu}_n)$, then it implies that $\mathcal{F}^n(\hat{\mathcal{W}}_n, \hat{\nu}_n)$ properly returned to $\mathcal{R}^*$, and $\mathcal{F}^n(\mathcal{W}_n, \nu_n) = \mathcal{F}^n(\hat{\mathcal{W}}_n, \hat{\nu}_n)|_{\mathcal{R}_*}$ is obtained by restricting the n-th image of its shadow family on $\mathcal{R}^*$. We also define a *stopping time* τ, a well as a return map T, such that for any $n \geq 0$, for ν-almost every $x \in \mathcal{W}_n$, we put

$$\tau(x) = n, \quad Tx := \mathcal{F}^{\tau(x)} x. \tag{5.1}$$

This also implies that for ν-a.e. x in the level set $(\tau = n)$, Tx has properly returned to $\mathcal{R}^*$ together with $\mathcal{F}^n \mathcal{W}_n$. Moreover, for any standard family $\mathcal{G} = (\mathcal{W}, \nu)$, we have the following decompositions:

$$\mathcal{G} = \sum_{n \geq 0} (\mathcal{W}_n, \nu_n), \quad T\mathcal{G} = \sum_{n \geq 0} (\mathcal{F}^n \mathcal{W}_n, \mathcal{F}^n_* \nu_n). \tag{5.2}$$

Here $T\mathcal{G}$ becomes a generalized standard family of index 0. This stopping time τ depends on the original family $\mathcal{G}$; and it is crucial in our coupling scheme. In addition according to our definition of T, we know that T only sees proper returns to $\mathcal{R}^*$. This is important in proving the Coupling Lemma 6.1.

5.3 Multiple returns to $\mathcal{R}^*$

For any $n \geq 1$, $m \geq n$, we define

$$\mathcal{D}_{n,m} = \{x \in \mathcal{R}^* \,:\, \tau(x) + \cdots + \tau(T^n x) = m\}. \tag{5.3}$$

One can check that $\mathcal{D}_{n,m} \subset \Gamma^s$ is an s-subset, such that $T^n \mathcal{D}_{n,m} = \mathcal{F}^m \mathcal{D}_{n,m}$. Clearly $T^n \mathcal{D}_{n,m}$ is a u-subset of $\mathcal{U}^*$.

Lemma 5.6. *Let*

$$\mathcal{G}^i = (\mathcal{W}^i, \nu^i) = ((W_\alpha, \nu_\alpha), \alpha \in \mathcal{A}^i, \lambda^i),$$

$i = 1, 2$, be two pseudo-generalized standard families of index zero, as defined in Definition 5.5, such that $\mathcal{W}^i$ properly cross $\mathcal{R}^*$. Assume that both ν^1 and ν^2 are probability measures. For any $n \geq 1$ and $k \geq n$, we have

$$|\nu^1(\mathcal{D}_{n,k}) - \nu^2(\mathcal{D}_{n,k})| \leq C_1 \nu^2(\mathcal{D}_{n,k}) \vartheta_1^n, \tag{5.4}$$

for some uniform constant $C_1 > 0$ and $\vartheta_1 \in (0, 1)$.

Proof. Let $\mathcal{G}^i = (\mathcal{W}^i, \nu^i) = ((W_\alpha, \nu_\alpha), \alpha \in \mathcal{A}^i, \lambda^i)$, $i = 1, 2$, be two pseudo generalized families of index 0. Then it follows from the definition that $T\mathcal{G}^i$ are both generalized standard families of index 0.

We fix any $n \geq 1$, for any $k \geq n$, let $U_{n,k}$ be the minimal collection of s-subsets containing $\mathcal{D}_{n,k}$, with

$$T^n U_{n,k} = \mathcal{F}^k U_{n,k}$$

being a u- subset properly crosses $\mathcal{U}^*$.

For any $\alpha \in \mathcal{A}^i$, let

$$W_{\alpha,k} = W_\alpha \cap U_{n,k},$$

and

$$\mathcal{W}^i_{n,k} := \{W_{\alpha,k}, \alpha \in \mathcal{A}^i\}.$$

Then we can start from the standard family

$$\mathcal{G}^i_{n,k} := (\mathcal{W}^i_{n,k}, \nu^i|_{U_{n,k}}).$$

Note that $T^n \mathcal{D}_{n,k}$ fully crosses $\mathcal{R}^*$ and

$$\begin{aligned}\int_{\alpha \in \mathcal{A}^i} \nu_\alpha(\mathcal{D}_{n,k})\, \lambda^i(d\alpha) &= \nu^i(\mathcal{D}_{n,k}) = T^n_* \nu^i(T^n \mathcal{D}_{n,k}) \\ &= T^n_*(\nu^i|_{U_{n,k}})(\mathcal{R}^*) = T^n_* \nu^i(T^n \mathcal{D}_{n,k}) \\ &= \int_{\alpha \in \mathcal{A}^i} T^n_* \nu_\alpha(T^n \mathcal{D}_{n,k})\, \lambda^i(d\alpha). \end{aligned} \tag{5.5}$$

It follows that the standard family $\mathcal{G}^i_{n,k}$ and its image $T^n \mathcal{G}^i_{n,k}$ have the same factor measures – mainly because unstable manifolds in $\mathcal{W}^i_{n,k}$ have not been cut by any singular curves under T^n.

(5.5) implies that it is enough to compare $T^n_*(\nu^1|_{U_{n,k}})$ and $T^n_*(\nu^2|_{U_{n,k}})$, which are associated with two standard families $T^n \mathcal{G}^i_{n,k}$ that fully returned to $\mathcal{U}^*$. For any $n \geq 1$, and $k \geq n$, we also denote

$$T^n \mathcal{G}^i_{n,k} = (\mathcal{V}^i_{n,k}, T^n_*(\nu^i|_{U_{n,k}})) = \{(T^n W_{\alpha,k}, T^n_* \nu_\alpha), \alpha \in \mathcal{A}^i, k \geq 1, \lambda^i\}.$$

Moreover, we denote

$$\mathcal{E}^i_{n,k} = \{(T^n W_{\alpha,k}, \mu_{\alpha,k}), \alpha \in \mathcal{A}^i, \lambda^i\},$$

as the associated μ- standard family of $T^n_* \nu_\alpha$ replaced by the u-SRB measure

$$\mu_{\alpha,k} := \mu_{T^n W_{\alpha,k}},$$

yet having the same index set as well as the same factor measure. We also denote

$$\mathcal{E}^i_{n,k} = (\mathcal{V}^i_{n,k}, \mu^i_{n,k})$$

for simplicity, such that for any measurable set $A \subset \mathcal{U}^*$,

$$\mu^i_{n,k}(A) = \int_{\alpha \in \mathcal{A}^i} \int_{x \in T^n W_{\alpha,k}} \chi_A(x)\, d\mu_{\alpha,k}\, \lambda^i(d\alpha). \tag{5.6}$$

Now Lemma 3.7 implies that $T^n \mathcal{G}^i_{n,k}$ can be approximated by its associated u-standard family $\mathcal{E}^i_{n,k}$ in the following sense:

$$|T^n_*(\nu^i|_{U_{n,k}})(\mathcal{R}^*) - \mu^i_{n,k}(\mathcal{R}^*)| \leq C C_F \mu^i_{n,k}(\mathcal{R}^*) \Lambda^{-n\gamma_0}. \tag{5.7}$$

Or we can write

$$T^n_*(\nu^i|_{U_{n,k}})(\mathcal{R}^*) = \mu^i_{n,k}(\mathcal{R}^*)(1 + \mathcal{O}(\Lambda^{-n\gamma_0})),$$

where $\mu^i_{n,k}$ was defined as in (5.6), and

$$|\mathcal{O}(\Lambda^{-n\gamma_0})| \leq C_1 \Lambda^{-n\gamma_0}.$$

For any $n \geq 1$, we denote $\mathcal{B}_n = \mathcal{B}_n(W)$ as the index set of $T^n W$. According to (3.24), we know that the pushforward image of a standard pair (W, μ_W) must be of the form $(T^n W, T^n_* \mu_W)$, where

$$T^n W = \{V_\beta, \beta \in \mathcal{B}_n\}$$

and for any measurable set $A \subset \mathcal{M}$,

$$T^n_* \mu_W(A) = \int_{\beta \in \mathcal{B}_n} \mu_\beta(A \cap V_\beta)\, \lambda^u_{\mathcal{M}}(d\alpha).$$

Thus the standard family $\mathcal{E}^1_{n,k}$ must come from the n-th iterations of $((W_{\alpha,k}, \mu_{W_{\alpha,k}}), \mathcal{A}^i, \lambda^i)$. This implies that

$$\mu^i_{n,k}(\mathcal{R}^*) = \int_{\alpha \in \mathcal{A}^i} \mu_{\alpha,k}(\mathcal{R}^*)\, \lambda^i(d\alpha). \tag{5.8}$$

Moreover, note that

$$\lambda^i(\mathcal{A}^i) = 1$$

as ν^i is a probability measure. So in order to compare $\mu^1_{n,k}(\mathcal{R}^*)$ and $\mu^2_{n,k}(\mathcal{R}^*)$, it is enough to compare $\mu_{\alpha,k}(\mathcal{R}^*)$ for $\alpha \in \mathcal{A}^1$ and $\alpha \in \mathcal{A}^2$.

Since for each $k \geq 1$, the stable boundaries of $U_{n,k}$ shrink exponentially fast under the iteration of T^n, thus the standard pairs in the family $\mathcal{E}^1_{n,k} = (\mathcal{V}^1_{n,k}, \mu^1_{n,k})$ should be very "close" to those in $\mathcal{E}^2_{n,k} = (\mathcal{V}^2_{n,k}, \mu^2_{n,k})$. Indeed note that unstable manifolds in $\mathcal{V}^1_{n,k}$ and those in $\mathcal{V}^2_{n,k}$ are at most $C\Lambda^{-n}$ far apart for some uniform constant depending only on the distortion constant $C_{\mathbf{r}}$ defined in (2.5). Now we can use the assumption (2.6) on distortion bounds for the Jacobian of the stable holonomy map defined by Γ^s. More precisely, for any $\alpha \in \mathcal{A}^1$, $\beta \in \mathcal{A}^2$, we define

$$\mathbf{h} : W_\alpha \to W_\beta$$

as the stable holonomy map, with

$$\mathbf{h}(x) = d\mu_\alpha / d\mu_\beta,$$

for any $x \in W_\alpha \cap \Gamma^s$ and

$$\mathbf{h}_n : T^n(W_\alpha \cap \Gamma_n) \to T^n(W_\beta \cap \Gamma_n).$$

Then by the absolute continuity property of the holonomy map, especially (2.6) and (2.7), we have

$$|\ln \mathbf{h}_n| \leq C_{\mathbf{r}} \vartheta_0^n,$$

as points in Γ_n have returned to M at least n times under the induced map T. Thus for any measurable collection of stable manifolds $A \subset \Gamma^s$, any standard pair $(T^n W_{\alpha,k}, \mu_{\alpha,k}) \in \mathcal{E}^1_{n,k}$ and $(T^n W_{\beta,k}, \mu_{\beta,k}) \in \mathcal{E}^2_{n,k}$, we have

$$\mu_{\alpha,k}(A \cap T^n W_{\alpha,k}) = \mu_{\beta,k}(A \cap T^n W_{\beta,k})(1 + \mathcal{O}(\vartheta_0^n)). \tag{5.9}$$

Since (5.9) is true for all $\alpha \in \mathcal{A}^1$ and all $\beta \in \mathcal{A}^2$, using the fact that $\lambda^i(\mathcal{A}^i) = 1$, we have by (5.8),

$$|\mu^1_{n,k}(\mathcal{R}^*) - \mu^2_{n,k}(\mathcal{R}^*)| \leq \mu^1_{n,k}(\mathcal{R}^*)(1 + C\vartheta_0^n).$$

Now combining with (5.7) and (5.5) , we get

$$|\nu^1(\mathcal{D}_{n,k}) - \nu^2(\mathcal{D}_{n,k})| \leq C_1 \nu^2(\mathcal{D}_{n,k}) \vartheta_1^n,$$

where

$$\vartheta_1 = \max\{\vartheta_0, \Lambda^{-\gamma_0}\}$$

and $C_1 > 0$ is a constant.

□

In the proof of our main results, we need the following lemma which enables us to define a fixed number of additional iterations after each step of coupling, to ensure that there are certain amount of measure properly returned to $\mathcal{R}^*$ simultaneously from both families.

Lemma 5.7. *For $\delta > 0$ as defined in Lemma 5.2, there exists $t_0 \geq 1$, with the following properties:*

(1) *For any pseudo-generalized standard family of index zero* $\mathcal{G}_0 = (\mathcal{W}_0, \nu_0)$, $\mathcal{F}_*^{t_0}\nu_0$ *has at least* δ *portion of measure* $\nu_0(\mathcal{M})$ *properly returned to* $\mathcal{R}^*$*;*

(2) *Let* $\mathcal{G}_m = (\mathcal{W}_m, \nu_m)$ *be a generalized standard family of index* m*, for* $m = 1, \cdots, t_0$*. We denote*

$$(\mathcal{W}, \nu) = \sum_{m=0}^{t_0} \mathcal{G}_m$$

and assume that ν *is a probability measure. Then the total amount of measure of* $(\mathcal{W}, \nu)$ *that properly returned to* $\mathcal{R}^*$ *under* $\mathcal{F}^{t_0}$*-iterations is at least* δ*.*

This is a very important lemma, and the proof is kind of lengthy, so we include the proof in Subsection 5.4.

5.4 Proof of Lemma 5.7

The following facts for renewal sequences were proved in a series of references, starting from Kolmogorov, [51, 43, 65], and mainly in [60].

Lemma 5.8. *Let* $\{p_k, k \geq 1\}$ *be a nonnegative sequence such that*

$$\sum_{k=1}^{\infty} p_k = 1.$$

Let $a_0 = 1$ *and for any* $n \geq 1$*, we set* $a_n = \sum_{k=1}^{n} p_k a_{n-k}$*. If*

$$\lambda := \sum_{k=1}^{\infty} k p_k < \infty,$$

and

$$\sum_{k=1}^{\infty} k^{\xi} p_k < \infty,$$

with some $\xi > 1$*, then*

$$a_n = \frac{1}{\lambda} + \frac{1}{\lambda^2} \sum_{k \geq n} \sum_{m=k+1}^{\infty} p_m + R_n(\xi),$$

where $R_n(\xi) = \mathcal{O}(n^{-\xi})$ *if* $\xi \geq 2$*; and* $R_n(\xi) = \mathcal{O}(n^{-2(\xi-1)})$ *if* $\xi \in (1, 2)$*.*

We define

$$p_k = \mu(\Gamma_k^s | \mathcal{R}^*)$$

for any $k \geq 1$. By Proposition 5.1, we let $\xi = \alpha$. Then we have $\mu(\tau^{\xi}) < \infty$. We first prove a lemma.

Lemma 5.9. *Let $(\mathcal{W},\nu)$ be a pseudo-generalized standard family with index zero, and $\nu(\mathcal{M})=1$. Then the total amount of measure of $(\mathcal{W},\nu)$ that properly returned to $\mathcal{R}^*$ after $\mathcal{F}^n$-iterations, which is denoted as r_n, for any $n\geq 1$, satisfies*

$$r_n=\frac{1}{\lambda}+\frac{1}{\lambda^2}P_n+R_n,$$

where we denote $1/\lambda=\mu(\mathcal{R}^*)$, $P_n=\sum_{i\geq n}\sum_{m=i+1}^{\infty}p_m$, and $R_n=\mathcal{O}(R_n(\xi))$.

Proof. Let $a_0=1$ and for any $n\geq 1$, we define

$$a_n=\sum_{k=1}^{n}p_k a_{n-k}.$$

Moreover, we denote

$$\lambda:=\sum_{k=1}^{\infty}kp_k=1/\mu(\mathcal{R}^*)<\infty,$$

and by the assumption that $\mu(\tau^\xi)<\infty$, we know that

$$\sum_{k=1}^{\infty}k^\xi p_k<\mu(\tau^\xi)<\infty.$$

Then the above renewal theory implies that

$$a_n=\frac{1}{\lambda}+\frac{1}{\lambda^2}\sum_{k\geq n}\sum_{m=k+1}^{\infty}p_m+R_n(\xi), \tag{5.10}$$

where $R_n(\xi)=\mathcal{O}(n^{-\xi})$ if $\xi\geq 2$; and $R_n(\xi)=\mathcal{O}(n^{-2(\xi-1)})$ if $\xi\in[1,2)$.

We define $\mathcal{D}_{k,n}$ as the set of all points in $\mathcal{R}^*$ that return to $\mathcal{R}^*$ exactly k times under $\mathcal{F}^n$, see (5.3). Then one can check that

$$\mathcal{D}_{k,n}=\{x\in\mathcal{R}^*\,:\,\tau(x)+\cdots+\tau(T^kx)=n\}.$$

Let A_n be the set of all points that returned to $\mathcal{R}^*$ under $\mathcal{F}^n$. Then

$$A_n=\{x\in\mathcal{R}^*\,:\,\tau(x)+\cdots+\tau(T^kx)=n,\ \text{for some } k=0,\cdots,n-1\}=\bigcup_{k=0}^{n-1}\mathcal{D}_{k,n},$$

as all points in $\mathcal{R}^*$ that will return to $\mathcal{R}^*$ after n-iterations. Let $a_0'=1$, and $a_n'=\mu(A_n|\mathcal{R}^*)$ as the total measure that will return to $\mathcal{R}^*$ properly. We define the measure $\hat{\mu}:=\mu|_{\mathcal{R}^*}/\mu(\mathcal{R}^*)$; clearly it is invariant under the return map to $\mathcal{R}^*$. Then we can check that

$$a_1'=\hat{\mu}(A_1)=\hat{\mu}(\Gamma_1^s)=a_1=p_1a_0.$$

Furthermore,

$$a_2' = \hat{\mu}(A_2) = \hat{\mu}(\mathcal{D}_{2,2}) + \hat{\mu}(\mathcal{D}_{1,2}) = \frac{\hat{\mu}(\Gamma_1^s \cap \mathcal{F}^{-1}\Gamma_1^s)}{\hat{\mu}(\Gamma_1^s)} p_1 + p_2 = \hat{\mu}(\mathcal{D}_{1,1}|\Gamma_1^s)p_1 + p_2 a_0.$$

For any $k \geq 1$, we define the conditional measure on Γ_k^s obtained by $\hat{\mu}$ as

$$\eta_k := \frac{\hat{\mu}|_{\Gamma_k^s}}{\hat{\mu}(\Gamma_k^s)}.$$

We can check that

$$\frac{\hat{\mu}(\Gamma_1^s \cap \mathcal{F}^{-1}\Gamma_1^s)}{\hat{\mu}(\Gamma_1^s)} = \mathcal{F}_*\eta_1(\Gamma_1^s).$$

Since both $\mathcal{F}_*\eta_1$ and $\hat{\mu}$ are probability measures with support properly cross $\mathcal{R}^*$, we can apply Lemma 5.6, to get

$$\mathcal{F}_*\eta_1(\Gamma_1^s) - \hat{\mu}(\Gamma_s^1) \leq C_1 \hat{\mu}(\Gamma_s^1)\vartheta_1,$$

which is equivalent as

$$|\hat{\mu}(\mathcal{D}_{2,2}) - a_1 p_1| \leq \varepsilon_2 := C_1 a_1 p_1 \vartheta_1.$$

Here $C_1 > 0$ and $\vartheta_1 \in (0,1)$ were given in lemma 5.6. Thus we get

$$a_2' = a_2 + \varepsilon_2.$$

Furthermore,

$$\begin{aligned}
a_3' &= \hat{\mu}(A_3) = \hat{\mu}(\mathcal{D}_{3,3}) + \hat{\mu}(\mathcal{D}_{2,3}) + \hat{\mu}(\mathcal{D}_{1,3}) \\
&= \hat{\mu}(\mathcal{F}^{-1}\Gamma_1^s \cap \mathcal{F}^{-2}\Gamma_1^s|\Gamma_1^s)p_1 + \hat{\mu}(\mathcal{F}^{-1}\Gamma_2^s|\Gamma_1^s)p_1 + \hat{\mu}(\mathcal{F}^{-2}\Gamma_1^s|\Gamma_2^s)p_2 + p_3 a_0 \\
&= \mathcal{F}_*\eta_1(\mathcal{D}_{2,2})p_1 + \mathcal{F}_*\eta_1(\Gamma_2^s)p_1 + \mathcal{F}_*^2\eta_2(\Gamma_1^s)p_2 + p_3 a_0 \\
&= \mathcal{F}_*\eta_1(A_2)p_1 + \mathcal{F}_*^2\eta_2(A_2)p_2 + p_3 a_0.
\end{aligned}$$

Using Lemma 5.6, we get

$$|\mathcal{F}_*\eta_1(A_2) - a_2'| \leq C_1 a_2' \vartheta_1,$$

as well as

$$|\mathcal{F}_*^2\eta_2(A_2) - a_1'| \leq C_1 a_1' \vartheta_1^2.$$

Combing above estimates, we get

$$a_3' = a_3 + \sum_{k=2}^{3} \varepsilon_k p_{3-k},$$

where we define $p_0 = 1$, and the error term ε_k satisfies

$$|\varepsilon_k| \leq C_1 \sum_{j=1}^{k-1} a_j' p_{k-j} \vartheta_1^{k-j}.$$

Inductively, assume for any $k \leq n-1$, we have

$$a'_k = a_k + \varepsilon_{k,k-1} + \varepsilon_{k,k-2}p_1 + \cdots + \varepsilon_2 p_{k-2},$$

with

$$|\varepsilon_m| \leq C_1 \sum_{j=1}^{m-1} p_{m-j} a'_j \vartheta_1^{m-j},$$

for any $m = 2, \cdots, k$.

Next we estimate a'_n. Note that

$$\begin{aligned} a'_n &= \hat{\mu}(A_n) \\ &= \mathcal{F}_* \eta_1(A_{n-1})p_1 + \mathcal{F}_*^2 \eta_2(A_{n-2})p_2 + \cdots + p_{n-1}\mathcal{F}_*^{n-1}\eta_{n-1}(A_1) + p_n a_0. \end{aligned}$$

Using Lemma 5.6, we get for any $k = 1, \cdots, n-1$,

$$|\mathcal{F}_*^k \eta_k(A_{n-k}) - a'_{n-k}| \leq C_1 a'_{n-k} \vartheta_1^k.$$

Combing above estimates as well as our assumptions, we get

$$a'_n = a_n + \sum_{k=2}^{n} \varepsilon_k p_{n-k}, \tag{5.11}$$

where again, we denote $p_0 = 1$, and the error term ε_n satisfies

$$|\varepsilon_n| \leq C_1 \sum_{j=1}^{n-1} a'_j p_{n-j} \vartheta_1^{n-j}.$$

Next we consider any any pseudo-generalized standard family of index zero, $(\mathcal{W}, \nu)$, with probability measure $\nu(\mathcal{R}^*) = 1$.

Then by Lemma 5.6, we can check that for any $k \geq 1$,

$$|\nu(\Gamma_k^s) - \hat{\mu}(\Gamma_k^s)| \leq C\hat{\mu}(\Gamma_s^k),$$

which implies that

$$|\nu(\Gamma_k^s) - p_k| \leq C p_k. \tag{5.12}$$

In particular, the total amount of measure in $\mathcal{F}((\mathcal{W}, \nu)$ that properly returned to $\mathcal{R}^*$ satisfies

$$r_1 := \nu(\Gamma_1^s) = \nu(\Gamma_1^s) = p_1 + \varepsilon_0 = a_1 + \varepsilon_0,$$

where $|\varepsilon_0| < Cp_1$.

We also define

$$r_0 := 1.$$

Let r_n be the total amount of measure in $\mathcal{F}^n(\mathcal{W},\nu)$ that properly returned to $\mathcal{R}^*$. Using Lemma 5.6, we first check r_2 satisfies

$$\begin{aligned} r_2 = \nu(\cup_{k=1}^2 \mathcal{D}_{k,2}) &= \sum_{k=1}^{2} \nu(\mathcal{D}_{k,2}) \\ &\leq \sum_{k=1}^{2} \hat{\mu}(\mathcal{D}_{k,2})(1 + C_1 \vartheta_1^k) \\ &= a_2' + C_1 \sum_{k=1}^{2} p_k a_{2-k}' \vartheta_1^k, \end{aligned}$$

where $C_1 > 0$ and $\vartheta \in (0,1)$ were uniform constants given in Lemma 5.6.

Inductively we can check that the total amount of measure in $\mathcal{F}^n(\mathcal{W},\nu)$ that properly returned to $\mathcal{R}^*$ satisfies

$$\begin{aligned} r_n &= \sum_{k=1}^{n} \nu(\mathcal{D}_{k,n}) \\ &\leq a_n' + C_1 \sum_{k=1}^{n} \hat{\mu}(\mathcal{D}_{k,n}) \vartheta_1^k \\ &= a_n' + C_1 \sum_{k=1}^{n} p_k a_{n-k}' \vartheta_1^k. \end{aligned}$$

Combining with (5.11), and using the fact that $a_n' \leq 1$ is bounded, we get

$$\begin{aligned} |r_n - a_n| &\leq 2C_1 \sum_{m=2}^{n} p_{n-m} \sum_{j=1}^{m} a_{m-j}' p_j \vartheta_1^j \\ &= 2C_1 \sum_{j=1}^{m} \left(\sum_{m=j}^{n} p_{n-m} a_{m-j} \right) p_j \vartheta_1^j \leq C p_n, \end{aligned}$$

for some uniform constant $C > 0$. According to Lemma 5.8, as well as (5.10), we know that

$$r_n = \frac{1}{\lambda} + \frac{1}{\lambda^2} P_n + R_n,$$

with $R_n = R_n(\xi) + C p_n = \mathcal{O}(R_n(\xi))$.

□

Next we consider the general case.

Lemma 5.10. *Let $\mathcal{G}_m = (\mathcal{W}_m, \nu_m)$ to be a generalized standard family of index m, for $m = 0, \cdots, t_0$. We denote*

$$(\mathcal{W},\nu) = \sum_{m=0}^{t_0} \mathcal{G}_m,$$

and assume ν is a probability measure. Then the total amount of measure of $(\mathcal{W},\nu)$ that properly returned to $\mathcal{R}^*$ after $\mathcal{F}^{t_0}$-iterations, which is denoted as s_{t_0}, satisfies

$$s_{t_0} = \lambda^{-1} + \sum_{k=0}^{t_0-1} \lambda^{-2} q_k (P_{t_0-k} + R^k_{t_0-k}) \geq \mu(\mathcal{R}^*),$$

where we denote for any $n \geq 1$, $P_n = \sum_{i\geq n}\sum_{m=i+1}^{\infty} p_m$, $q_n = \nu_n(\mathcal{M})$, and $R^k_n = O(R_n(\xi))$.

Proof. Let

$$(\mathcal{W},\nu) = \sum_{m=0}^{t_0} \mathcal{G}_m$$

be the sum of t_0 generalized standard families, with probability measure ν. We define

$$q_m = \mathcal{F}^m_* \nu_m(\mathcal{R}^*).$$

Note that once $\mathcal{F}^m\mathcal{G}_m$ arrives at $\mathcal{R}^*$, its further iterations will return to $\mathcal{R}^*$ according to the Markov decomposition

$$\Gamma^s = \cup_{k\geq 1}\Gamma^s_k.$$

Note that $(\mathcal{W}_0,\nu_0)$ has index 0, we denote $r^0_0 = q_0$ and $p_0 = 0$. Then we can apply the above lemma to get that the total amount of measure in $\mathcal{F}^n(\mathcal{W}_0,\nu_0)$ that properly returned to $\mathcal{R}^*$ satisfies

$$r^0_n = q_0\left(\frac{1}{\lambda} + \frac{1}{\lambda^2}(P_n + R^0_n)\right),$$

where $R^0_n = \mathcal{O}(R_n(\xi))$.

Inductively, one can show that in the generalized family $(\mathcal{W}_k,\nu_k)$, for $k = 1,\cdots,n$, the total amount of measure in $\mathcal{F}^n(\mathcal{W}_k,\nu_k)$ that properly returned to $\mathcal{R}^*$ satisfies

$$r^k_n = q_k\left(\frac{1}{\lambda} + \frac{1}{\lambda^2}(P_{n-k} + R^k_{n-k})\right).$$

Thus the total amount of measure of $(\mathcal{W},\nu)$ that properly returned to $\mathcal{R}^*$ after $\mathcal{F}^n$-iterations, which is denoted as s_n, for any $n \geq 0$, satisfies

$$\begin{aligned} s_n &= r^0_n + r^1_n + \cdots + r^n_n \\ &= \lambda^{-1}\sum_{k=0}^{n} q_k + \sum_{k=0}^{n}\lambda^{-2} q_k (P_{n-k} + R^k_{n-k}). \end{aligned}$$

Since P_n is a decreasing sequence, thus we can check that there exists $C > 0$ such that

$$\sum_{k=0}^{n} q_k P_{n-k} \leq C\max\{P_n, q_n\}.$$

Thus for any $n < t_0$, we have

$$\begin{aligned} s_n &= \lambda^{-1} \sum_{k=0}^{n} q_k + \mathcal{O}(P_n + q_n) \\ &= \lambda^{-1}(1 - \sum_{k=n+1}^{\infty} q_k) + \mathcal{O}(P_n + q_n) \\ &= \lambda^{-1} - \lambda^{-1} \sum_{k=n+1}^{t_0} q_k + \mathcal{O}(P_n + q_n). \end{aligned}$$

Moreover, for $n = t_0$, we have

$$s_{t_0} = \lambda^{-1} + \sum_{k=0}^{t_0} \lambda^{-2} q_k (P_{t_0-k} + R^k_{t_0,k}) \geq \lambda^{-1}.$$

Since δ can be chosen such that $\lambda^{-1} = \mu(\mathcal{R}^*) > \delta$, we have proved the claimed results. □

6 Coupling Lemma for the original system

In this section, we will prove the Coupling Lemma for the original nonuniformly hyperbolic map, which is new to our knowledge, as the construction is significantly different from that for systems with uniformly hyperbolicity. This will enable us to define the coupling decompositions of probability measures on $\mathcal{M}$, which will be used to investigate the rate of decay of correlations of iterations of those measures.

6.1 Statement of the Coupling Lemma

We now state the coupling lemma for the original nonuniformly hyperbolic system $(\mathcal{F}, \mu)$.

Lemma 6.1. *Let $\mathcal{G}^i := (\mathcal{W}^i, \nu^i)$ be two standard families, with $d\nu^i = g^i d\mu$, $i = 1, 2$, where $g^i \in \mathcal{H}^+(\gamma)$ are probability density functions with $\gamma \geq \gamma_0$.*
(C1) *There exist $N_1 = N_1(\nu_1, \nu_2) \geq 1$, and uniform constants $C_0, C_1, C > 0$, such that for any $n \geq 1$, there is a decomposition*

$$\mathcal{F}^{N_1} \mathcal{G}^i = \sum_{k=1}^{n} (\mathcal{W}^i_k, \nu^i_k) + (\bar{\mathcal{W}}^i_n, \bar{\nu}^i_n),$$

for $i = 1, 2$, with the following properties for any $k = 1, \cdots, n$:

(i) $(\mathcal{W}^i_k, \nu^i_k)$ is a generalized standard family with index k;

(ii) For any measurable function f that is constant on each $W^s \in \Gamma^s$, we have $\mathcal{F}^k_* \nu^1_k(f) = \mathcal{F}^k_* \nu^2_k(f)$;

(iii) For any $f \in \mathcal{H}^-(\gamma_0)$, the uncoupled measure $\bar{\nu}_k^i$ satisfies:

$$|\bar{\nu}_k^i(f)| \le C_0\|f\|_\infty \nu^i(C_{k,b} \cup (R > k)) + C_1\|g^i\|_{C^\gamma}^+ k^{-\alpha_0} \le C\|f\|_\infty \max\{\|g^1\|_{C^\gamma}^+, \|g^2\|_{C^\gamma}^+\} k^{1-\alpha_0};$$

(iv) Moreover, $\bar{\nu}_k^i(\mathcal{M}) = \mathcal{F}_*^{N_1}\nu^i(\tau \ge k) + \mathcal{O}(\nu^i(k \le \tau \le k + t_0))$.

(i) If both standard families $\mathcal{G}^1$ and $\mathcal{G}^2$ are proper family, then N_1 is a uniform constant.

*(**C2**) Moreover, there exists $C_2 > 0$, such that for any k large, the portion of measure coupled at k satisfies:*

$$\mathcal{F}_*^k \nu_k^i(\mathcal{R}^*) \le C_2 k^{1-\alpha_0+\varepsilon_1}.$$

The proof of this Coupling Lemma can be found in Subsection 6.3, after we describe in detail the coupling procedure for measures that properly returned to $\mathcal{R}^*$ in next subsection. Note that item (**C2**) looks very artificial, however it can be easily verified because of property **(C1)** and assumption (**H1**). The main reason for such a condition is that we need to compare the measure $\mathcal{F}_*^n \nu_n^i(\mathcal{R}^*)$ with $n^{-\alpha_0}$ later in the proof of Theorem 1, so just for convenience we add this property here.

We begin by considering a special situation, i.e. $\mathcal{G}^i = (\mathcal{W}^i, \nu^i)$, $i = 1, 2$, are two generalized standard families with index 0, and $\nu^1(\mathcal{R}^*) = \nu^2(\mathcal{R}^*)$. We first prove a lemma describing the coupling process which will be used in our proof of Lemma 6.1.

Lemma 6.2. *Assume that for $i = 1, 2$, $\mathcal{G}^i = (\mathcal{W}^i, \nu^i)$ are generalized standard families with index 0, and*

$$\min\{\nu^1(\Gamma^s), \nu^2(\Gamma^s)\} > 0.$$

Then there exist a generalized standard family $\mathcal{E}^i = (\mathcal{W}^i, \eta^i)$ with index 0, and

$$\mathcal{K}^i := \mathcal{G}^i - \mathcal{E}^i = (\mathcal{W}^i, \xi^i),$$

with the following properties:
*(**a**) $\mathcal{E}^1$ and $\mathcal{E}^2$ are coupled in the following sense:*
*(**a1**) For any $f \in \mathcal{H}^-(\gamma_0)$ that is constant on each $W^s \in \Gamma^s$, we have $\eta^1(f) = \eta^2(f)$;*
*(**a2**) The total coupled measure satisfies*

$$d := \eta^i(\Gamma^s) = c_0 \min\{\nu^1(\Gamma^s), \nu^2(\Gamma^s)\},$$

where $c_0 \in [1/3, 1/2]$.
*(**b**) The remaining uncoupled family $\mathcal{K}^i$ is a pseudo-generalized family with index 0, i.e. it has the property that $\mathcal{F}^n(\mathcal{K}^i|_{\Gamma_n^s})$ becomes a generalized standard family of index 0, for any $n \ge 1$.*
*(**c**) For any measurable collection of unstable manifolds $A \subset \Gamma^s$, the remaining uncoupled measure can be calculated as:*

$$|\nu^1(A) - \nu^2(A)| = |\xi^1(A) - \xi^2(A)|.$$

Proof. Since for $i = 1, 2$, $\mathcal{G}^i$ is a generalized standard family of index 0, by definition, $\mathcal{G}^i$ has a shadow family, which is a proper family that fully crosses $\mathcal{R}^*$ and has density function $g^i \in \mathcal{H}^+(\gamma)$, with $\gamma \geq \gamma_0$ and such that $g^i|_{\mathcal{R}^*}$ is the density function of ν^i. Then we denote

$$\mathcal{G}^i = \{(W_\alpha, \nu_\alpha) \,:\, \alpha \in \mathcal{A}^i, \lambda^i\},$$

such that for any measurable set $A \subset \mathcal{M}$,

$$\nu^i(A \cap \Gamma^s) = \int_{\alpha \in \mathcal{A}^i} \int_{W_\alpha \cap A} g^i_\alpha \, d\mu_\alpha \, \lambda^i(d\alpha),$$

where

$$g^i_\alpha = g^i / \mu_\alpha(g^i)$$

with

$$d\nu^i_\alpha = g^i_\alpha \, d\mu_\alpha,$$

and

$$\lambda^i(d\alpha) = \mu_\alpha(g^i)\lambda^u(d\alpha).$$

Clearly,

$$\lambda^i(\mathcal{A}^i) = \nu^i(\Gamma^s) = \mu(g^i|_{\mathcal{R}^*}).$$

For any $\alpha \in \mathcal{A}^i$, as the standard pair (W_α, ν^i_α) properly crosses Γ^s, by (3.14), we know that the density function satisfies

$$g^i_\alpha \geq e^{-0.01} \geq 0.9.$$

Thus one should be able to match at least 1/2 portion of measures from both families along stable manifolds in Γ^s. More precisely, we take $c_0 \in [1/3, 1/2]$, as we have flexibility to choose a Hölder function $\rho^i_\alpha \in \mathcal{H}^+(\gamma_0)$, such that

$$\rho^i_\alpha \in (g^i_\alpha/3, g^i_\alpha/2),$$

for any $\alpha \in \mathcal{A}^i$, such that

$$\int_{\alpha \in \mathcal{A}^i} \mu_\alpha(\rho^i_\alpha \cdot \mathbf{I}_{\mathcal{R}^* \cap W_\alpha}) \, d\lambda^i(\alpha) = c_0 \min\{\nu^1(\Gamma^s), \nu^2(\Gamma^s)\}, \quad \forall i = 1, 2. \tag{6.1}$$

Now we are ready to define the coupled families $\mathcal{E}^i$ for $i = 1, 2$. (6.1) implies that for any $\alpha \in \mathcal{A}^i$, one can define a standard pair corresponding to the measure defined by the density function ρ^i_α on W_α, denoted as (W_α, η_α). Let

$$\mathcal{E}^i := (\mathcal{W}^i, \eta^i) = ((W_\alpha, \eta_\alpha)|_{\Gamma^s}, \lambda^i)$$

be the corresponding generalized standard family with index zero. One could choose ρ^i_α and c_0 carefully to make sure that for any measurable collection A of stable manifolds in Γ^s, we have

$$\eta^1(A) = \eta^2(A),$$

and

$$d := \eta^i(\Gamma^s \cap \mathcal{W}^i) = c_0 \min\{\nu^1(\Gamma^s), \nu^2(\Gamma^s)\}.$$

This verifies items **(a1)-(a2)**.

Next, we define the remaining uncoupled family $\mathcal{K}^i$ by subtracting the density of η^i from ν^i. More precisely, for any $\alpha \in \mathcal{A}^i$, we subtract ρ^i_α from the density function g^i_α. The remaining family $\mathcal{G}^i \setminus \mathcal{E}^i$ is denoted as $\mathcal{K}^i$, which may not have the required regularity of being a generalized family. We apply Lemma 4.3 and using the definition of n_1 as well as Lemma 5.1, it follows that restricted on Γ^s_n, the family $\mathcal{F}^n(\mathcal{K}^i|_{\Gamma^s_n})$ becomes a generalized standard family with index 0, for any $n \geq 1$. Note that (**a2**) implies that, for any measurable collection of stable manifolds $A \subset \Gamma^s$,

$$\nu^1(A) - \nu^2(A) = \eta^1(A) - \eta^2(A) + \xi^1(A) - \xi^2(A) = \xi^1(A) - \xi^2(A). \tag{6.2}$$

Combining above facts, we get

$$\mathcal{G}^i = \mathcal{E}^i + \mathcal{K}^i$$

satisfying **(a)-(c)**, as claimed. □

6.2 Outline of the proof of the Coupling Lemma

Let $\mathcal{G}^i := (\mathcal{W}^i, \nu^i)$ be two standard families for $i = 1, 2$. We first describe an outline of the coupling scheme using above lemmas. The main difficulty of the coupling process is to guarantee frequently, simultaneous returns of certain portion of the conditional measure from both families. This fact is not straightforward because of the nonuniformly hyperbolicity. We overcome this difficulty by using the Markov structure of the generalized Young Tower based on $\mathcal{R}^* = \cup_{m\geq 1}\Gamma^s_m$ as introduced by Proposition 5.1, as well as the renewal theory proved by [43, 51, 60]. More precisely, we follow the 3 steps below:

(I) We first define $N_1 \geq 1$, such that both $\mathcal{F}^{N_1}_* \nu^1$ and $\mathcal{F}^{N_1}_* \nu^2$ have at least δ portion of measure properly returned to $\mathcal{R}^*$. We further decompose

$$\mathcal{F}^{N_1}\mathcal{G}^i = \sum_{m=0}^{\infty} \mathcal{G}^i_{1,m},$$

such that $\mathcal{F}^m\mathcal{G}^i_{1,m}$ is a generalized standard family of index 0.

(II) Now we have that both $\mathcal{G}^1_{1,0}$ and $\mathcal{G}^2_{1,0}$ are generalized standard family of index 0. By applying Lemma 6.2 we can couple at least $\delta/3$ of measures from both families, to get

$$\mathcal{G}^i_{1,0} = \mathcal{E}^i_1 + \mathcal{K}^i_1,$$

where $\mathcal{E}^i_1$ is the coupled family at this step. Since $\mathcal{K}^i_1$ is a pseudo-generalized standard family of index 0, we apply Lemma 5.7, which guaranteed that

at least δ portion of measure in this family will return to $\mathcal{R}^*$ after another t_0 iterations.

Thus the remaining families are split into two parts, $\hat{\mathcal{K}}_1^i$ and

$$\bar{\mathcal{G}}_1^i = \sum_{m \geq t_0+1} \mathcal{G}_{1,m}^i,$$

where $\hat{\mathcal{K}}_1^i$ consists of the remaining family $\mathcal{K}_1^i$ as well as $\mathcal{G}_{1,m}^i$ for $m = 1, \cdots, t_0$.

(III) Now we are in the similar situation as the beginning of step (II). We evolve both families under $\mathcal{F}^{t_0}$, and consider the higher iteration map $\mathcal{F}^{t_0}$. By applying Lemma 5.7, it is guaranteed that $\mathcal{F}^{t_0}\hat{\mathcal{K}}_1^i$ has at least $\delta > 0$ portion of the measure properly returned to $\mathcal{R}^*$. We again apply Lemma 6.2 to couple at least $\delta/3$ portion of part of the remaining measure from each family. The above algorithm is recurrent, thus by induction, at least the same portion $\delta/3$ of remaining probability measures could be coupled by applying Lemma 6.2 under every iterations t_0.

An intuitive way of visualizing this coupling scheme is that, although both families ν^1 and ν^2 may have totally different first return distributions to $\mathcal{R}^*$, once their supports properly crosses $\mathcal{R}^*$ for the first time, the returning measures will have almost "uniform" mixing, according to the Markov structure of $\mathcal{R}^*$. The processes can be approximated by a renewal process associated with an ergodic Markov chain with countably many states.

6.3 Proof of the Coupling Lemma

Step 0. Evolving both families under a large N_1 iterations.

Let $\mathcal{G}^i := (\mathcal{W}^i, \nu^i)$ be two standard families, with ν^i being probability measure for $i = 1, 2$. By Lemma 5.2, there exists N_1, such that for any $n \geq N_1$, both $\mathcal{F}_*^n \nu^1$ and $\mathcal{F}_*^n \nu^2$ are proper families and all have at least δ portion of measure properly returned to $\mathcal{R}^*$, where δ was given in Proposition 5.1. Since we have freedom to pick N_1 large enough, we use the same δ as in Proposition 5.1 just for convenience. Moreover, we choose $N_1 > t_0$, as t_0 is a uniform constant. This also implies that if both families $\mathcal{G}^1$ and $\mathcal{G}^2$ are proper, then N_1 can be chosen to be uniform (for all proper families). Then we consider

$$\mathcal{G}_1^i := \mathcal{F}^{N_1}\mathcal{G}^i = (\mathcal{W}_1^i, \nu_1^i).$$

We decompose

$$\mathcal{G}_1^i = \sum_{m=0}^{\infty} \mathcal{G}_{1,m}^i = \sum_{m=0}^{\infty} (\mathcal{W}_{1,m}^i, \nu_{1,m}^i),$$

according to (5.2), where $\mathcal{G}_{1,m}^i$ is a generalized standard family with index m, such that $\mathcal{F}^m \mathcal{W}_{1,m}^i$ properly returns $\mathcal{R}^*$.

Since below we will couple measures every t_0-iterations, we first split the above sum into blocks according to the number of steps, such that for any $k \geq 1$:

$$Q_k^i = \sum_{m=(k-1)t_0+1}^{kt_0} \nu_{1,m}^i(\mathcal{M}), \qquad Q_0^i = \nu_{1,0}^i(\mathcal{M}). \tag{6.3}$$

Here Q_k^i is the total measure in ν_1^i that have firstly returned to $\mathcal{R}^*$ between time interval $[(k-1)t_0+1, kt_0]$.

Moreover, by our choice of N_1, we know that

$$\nu_1^i(\Gamma^s) = \nu_{1,0}^i(\Gamma^s) > \delta, \tag{6.4}$$

which is important for our estimation of the coupling speed.

Step 1. Capture and then coupling along Γ^s.

Since the family $\mathcal{G}_{1,0}^i$ is a generalized standard family with index 0, we can apply Lemma 6.2, to get a decomposition

$$\mathcal{G}_{1,0}^i = \mathcal{E}_1^i + \mathcal{K}_1^i,$$

with

$$\mathcal{E}_1^i = (\mathcal{W}_{1,0}^i, \eta_1^i),$$

and

$$\mathcal{K}_1^i = (\mathcal{W}_{1,0}^i, \xi_1^i)$$

where $\mathcal{E}_1^i$ is a generalized standard family of index 0. Note that $\mathcal{E}_1^1$ and $\mathcal{E}_1^2$ are coupled along stable manifolds in Γ^s. More precisely, for any $f \in \mathcal{H}^-(\gamma_0)$ that is constant on each $W^s \in \Gamma^s$,

$$\eta_1^1(f) = \eta_1^2(f) \tag{6.5}$$

Moreover,

$$\eta_1^i(\Gamma^s) = c_0 \min\{\nu_{1,0}^1(\Gamma^s), \nu_{1,0}^2(\Gamma^s)\},$$

where c_0 is chosen in $[1/3, 1/2]$. Thus by (6.4), we have

$$d_1^i := \frac{\eta_1^i(\Gamma^s)}{\nu_{1,0}^i(\Gamma^s)} \geq \frac{\delta}{3}.$$

This implies that we can "match" at least 1/3 portion of each measure from $\nu_{1,0}^1$ and $\nu_{1,0}^2$, as they both have proper returned to $\mathcal{R}^*$.

By Lemma 5.7, we know that $\mathcal{F}^{t_0}\mathcal{K}_1^i$ has at least δ portion of measure properly returned to $\mathcal{R}^*$. We denote

$$s_1^i = \xi_1^i(\mathcal{R}^*) + Q_1^i = \nu_{1,0}^i(\Gamma^s)(1-d_1) + Q_1^i.$$

We put

$$\hat{\mathcal{K}}_1^i = (\hat{\mathcal{W}}_1^i, \hat{\nu}_1^i) := \frac{1}{s_1^i}\left(\mathcal{K}_1^i + \sum_{m=1}^{t_0} \mathcal{G}_{1,m}^i\right),$$

and the total remaining family is denoted as

$$\tilde{\mathcal{G}}_1^i := \sum_{m \geq t_0+1} \mathcal{G}_{1,m}^i.$$

Now we have shown that the total uncoupled portion after ther first step can be represented as:

$$\mathcal{G}_1^i - \mathcal{E}_1^i = s_1^i \hat{\mathcal{K}}_1^i + \tilde{\mathcal{G}}_1^i. \tag{6.6}$$

To summarize, the remaining families contains two parts:

(1a) $\tilde{\mathcal{G}}_1^i$ is the part that has not yet reached $\mathcal{R}^*$;

(1b) $s_1^i \hat{\mathcal{K}}_1^i$ corresponds to leftover in the portion that has reached $\mathcal{R}^*$ at the first step.

Step 2. Release and recapture.

Note that $\hat{\mathcal{K}}_1^i$ has a probability measure, for $i = 1, 2$. We claim that both $\mathcal{F}^{t_0}\hat{\mathcal{K}}_1^1$ and $\mathcal{F}^{t_0}\hat{\mathcal{K}}_1^2$ have at least δ portion of measures properly returned to $\mathcal{R}^*$. Clearly, this follows from Lemma 5.7.

Since we are in a similar situation as in step 1, we can apply Lemma 6.2, to get a decomposition

$$\mathcal{F}^{t_0}\hat{\mathcal{K}}_1^i = \mathcal{E}_2^i + \mathcal{K}_2^i + \mathcal{U}_2^i,$$

with

$$\mathcal{E}_2^i = (\mathcal{W}_2^i, \eta_2^i), \quad \mathcal{K}_2^i = (\mathcal{W}_2^i, \xi_2^i).$$

Here $\mathcal{U}_2^i$ denotes those of $\mathcal{F}^{t_0}\hat{\mathcal{K}}_1^i$ that have not arrived at $\mathcal{R}^*$. Note that $\mathcal{E}_2^1$ and $\mathcal{E}_2^2$ are coupled along stable manifolds in Γ^s. More precisely, for any measurable collection of stable manifolds we have: $A \subset \Gamma^s$,

$$\eta_2^1(A) = \eta_2^2(A) \tag{6.7}$$

Moreover, we denote

$$d_2^i := \eta_2^i(\mathcal{R}^*)/\mathcal{F}_*^{t_0}\hat{\nu}_1^i(\mathcal{R}^*) \geq \frac{\delta}{3}.$$

We apply the first statement of Lemma 5.7 to $\mathcal{K}_2^1$ and $\mathcal{K}_2^2$, as they are pseudo -generalized standard families of index zero. Thus both $\mathcal{F}^{t_0}\mathcal{K}_2^1$ and $\mathcal{F}^{t_0}\mathcal{K}_2^2$ have at least δ portion of measures properly returned to $\mathcal{R}^*$. Let

$$s_2^i = s_1^i(1 - d_2^i) + Q_2^i,$$

and we define

$$\hat{\mathcal{K}}_2^i = (\hat{\mathcal{W}}_2^i, \hat{\nu}_2^i) := (s_1^i \mathcal{K}_2^i + s_1^i \mathcal{U}_2^i + \sum_{m=t_0+1}^{2t_0} \mathcal{F}^{t_0} \mathcal{G}_{1,m}^i)/s_2^i.$$

Thus the total remaining family can be represented as

$$\bar{\mathcal{G}}_2^i := \sum_{m \geq 2t_0+1} \mathcal{F}^{t_0} \mathcal{G}_{1,m}^i.$$

Combining with (6.6), we know that the total uncoupled portion at the second step can be represented as:

$$\mathcal{F}^{t_0} \mathcal{G}_1^i - \mathcal{F}^{t_0} \mathcal{E}_1^i - s_1^i \mathcal{E}_2^i = s_2^i \hat{\mathcal{K}}_2^i + \tilde{\mathcal{G}}_2^i. \tag{6.8}$$

According to lemma 5.10 and Lemma 5.7, both $\mathcal{F}^{t_0} \hat{\mathcal{K}}_2^1$ and $\mathcal{F}^{t_0} \hat{\mathcal{K}}_2^2$ have at least δ portion of measures properly returned to $\mathcal{R}^*$. Thus we are in similar situation as in the beginning of step 2.

Step 3. Coupling at the repeated proper returns.

Next we consider higher iterations by induction, with $k \geq 2$. We assume s_j^i is well defined, for $j = 1, \cdots, k-1$, and let

$$L := \mathcal{F}^{t_0}.$$

We assume for $k \geq 1$,

$$L^{k-1} \mathcal{G}_1^i = L^{k-1} \mathcal{E}_1^i + s_1^i L^{k-2} \mathcal{E}_2^i + \cdots + s_{k-1}^i L \mathcal{E}_k^i + s_k^i \hat{\mathcal{K}}_k^i + \tilde{\mathcal{G}}_k^i, \tag{6.9}$$

where the coupled family at the k-th step is denoted as

$$\mathcal{E}_k^i = (\mathcal{W}_k^i, \eta_k^i),$$

with

$$d_k^i \geq \delta/3,$$

and

$$s_k^i := s_{k-1}^i (1 - d_k^i) + Q_k^i. \tag{6.10}$$

Moreover, we put

$$\hat{\mathcal{K}}_k^i = (\hat{\mathcal{W}}_k^i, \hat{\nu}_k^i) := \left(s_{k-1}^i \mathcal{K}_k^i + s_{k-1}^i \mathcal{U}_k^i + \sum_{m=(k-1)t_0+1}^{kt_0} L^{k-1} \mathcal{G}_{1,m}^i \right) / s_k^i,$$

The remaining family in $\mathcal{F}_*^{kt_0} \mathcal{G}_1^i$ have yet not reached $\mathcal{R}^*$ is denoted as:

$$\tilde{\mathcal{G}}_k^i := \sum_{m \geq kt_0+1} \mathcal{F}^{(k-1)t_0} \mathcal{G}_{1,m}^i. \tag{6.11}$$

To summarize, at the k-th step of the coupling process, the candidate of generalized families in $\mathcal{F}^{kt_0}\mathcal{G}_1^i$ that we will perform the coupling procedure is denoted as

$$s_k^i\hat{\mathcal{K}}_k^i := s_{k-1}^i(\mathcal{K}_k^i + \mathcal{U}_k^i) + \left(\mathcal{F}^{t_0}\tilde{\mathcal{G}}_{k-1}^i - \tilde{\mathcal{G}}_k^i\right).$$

It comes from two source of measures:

(a) One is the newly arrived measure between these $k-1$ and k-th step, i.e.

$$Q_k^i = \mathcal{F}^{t_0}\tilde{\mathcal{G}}_{k-1}^i(\mathcal{M}) - \tilde{\mathcal{G}}_k^i(\mathcal{M});$$

(b) The other part comes from the leftover of the $k-1$-th step coupling – i.e. those measures have properly returned to $\mathcal{R}^*$ at the $k-1$-th step, yet not being coupled, which we denote as

$$\mathcal{K}_k^i + \mathcal{U}_k^i := \mathcal{F}^{t_0}\hat{\mathcal{K}}_{k-1}^i - \mathcal{E}_k^i,$$

see below (6.14).

We apply the second statement of Lemma 5.7 to $\hat{\mathcal{K}}_k^1$ and $\hat{\mathcal{K}}_k^2$, then we know that both $\mathcal{F}^{t_0}\hat{\nu}_k^i$ have at least δ portion properly returned to $\mathcal{R}^*$.

Thus we can apply Lemma 6.2, to get a decomposition

$$L\hat{\mathcal{K}}_k^i = \mathcal{E}_{k+1}^i + \mathcal{K}_{k+1}^i + \mathcal{U}_{k+1}^i,$$

with

$$\mathcal{E}_{k+1}^i = (\mathcal{W}_{k+1}^i, \eta_{k+1}^i)$$

and

$$\mathcal{K}_{k+1}^i = (\mathcal{W}_{k+1}^i, \xi_{k+1}^i)$$

being a pseudo-generalized standard family of index 0; and $\mathcal{U}_{k+1}^i$ being those of $L\hat{\mathcal{K}}_k^i$ that have not arrived at $\mathcal{R}^*$ at this moment. Note that $\mathcal{E}_{k+1}^1$ and $\mathcal{E}_{k+1}^2$ are coupled along stable manifolds in Γ^s. More precisely, it implies that for any measurable collection of stable manifolds $A \subset \Gamma^s$,

$$\eta_{k+1}^1(A) = \eta_{k+1}^2(A) \tag{6.12}$$

Moreover,

$$d_{k+1}^i := \eta_{k+1}^i(\mathcal{R}^*)/L_*\hat{\nu}_k^i(\mathcal{R}^*) \geq \frac{\delta}{3}.$$

We now define

$$s_{k+1}^i = s_k^i(1 - d_{k+1}^i) + Q_{k+1}^i. \tag{6.13}$$

Next we denote

$$\hat{\mathcal{K}}_{k+1}^i = (\hat{\mathcal{W}}_{k+1}^i, \hat{\nu}_{k+1}^i) := \left(s_k^i\mathcal{K}_{k+1}^i + s_k^i\mathcal{U}_{k+1}^i + \sum_{m=kt_0+1}^{(k+1)t_0} L^k\mathcal{G}_{1,m}^i\right)/s_{k+1}^i.$$

and the total remaining family as

$$\tilde{\mathcal{G}}^i_{k+1} := (\tilde{\mathcal{W}}^i_{k+1}, \tilde{\nu}^i_{k+1}) := \sum_{m \geq (k+1)t_0+1} L^k \mathcal{G}^i_{1,m}.$$

Combining with (6.9), we have shown that

$$L^k \mathcal{G}^i_1 = L^k \mathcal{E}^i_1 + s^i_1 L^{k-1} \mathcal{E}^i_2 + \cdots + s^i_k L \mathcal{E}^i_{k+1} + s^i_{k+1} \hat{\mathcal{K}}^i_{k+1} + \tilde{\mathcal{G}}^i_{k+1}. \tag{6.14}$$

Inductively, we can also get a formula for s^i_{k+1}:

$$\begin{aligned} s^i_{k+1} &= Q^i_{k+1} + (1-d^i_{k+1})Q^i_k + (1-d^i_{k+1})(1-d^i_k)Q^i_{k-1} + \cdots + \Pi^{k+1}_{j=0}(1-d^i_j)Q^i_0 \\ &\leq \sum_{m=0}^{k+1} Q^i_{k+1-m}\zeta^m \leq CQ^i_{k+1}, \end{aligned}$$

where $\zeta = 1 - \delta/3$.

Thus the total uncoupled measure at $k+1$-th step is:

$$\hat{\nu}^i_{k+1}(\mathcal{M}) + \tilde{\nu}^i_{k+1}(\mathcal{M}) = s^i_{k+1} + \tilde{\nu}^i_{k+1}(\mathcal{M}) = \tilde{\nu}^i_{k+1}(\mathcal{M}) + \mathcal{O}(Q^i_{k+1}) \tag{6.15}$$

$$= \nu^i_1(\tau > kt_0) + \mathcal{O}(\nu^i_1(kt_0 \leq \tau \leq (k+1)t_0)). \tag{6.16}$$

Step 4. Rearrange the coupled measures by the real iteration time under $\mathcal{F}$.

Now we rearrange the above coupled and uncoupled families according to the real iteration time under $\mathcal{F}$. Then we have shown that there exist $N_1 = N_1(\nu^1, \nu^2) \geq 1$, such that for any $n \geq 1$, there is a decomposition

$$\mathcal{F}^{N_1}\mathcal{G}^i = \sum_{k=0}^{n} (\mathcal{W}^i_k, \nu^i_k) + (\bar{\mathcal{W}}^i_n, \bar{\nu}^i_n),$$

for $i = 1, 2$. Here for any $n \geq 0$, $(\mathcal{W}^i_n, \nu^i_n)$ is a generalized standard family with index n, which is defined such that $(\mathcal{W}^i_n, \nu^i_n)$ is empty if $n/t_0 \neq [k/t_0]$; and

$$\mathcal{F}^n(\mathcal{W}^i_n, \nu^i_n) := s^i_{n/t_0} \mathcal{E}^i_{n/t_0+1},$$

otherwise. Here we also define $s^i_0 = 1$. Moreover, the uncoupled measure is defined such that

$$\mathcal{F}^n(\bar{\mathcal{W}}^i_n, \bar{\nu}^i_n) := s^i_{[n/t_0]+1} \hat{\mathcal{K}}^i_{[n/t_0]+1} + \tilde{\mathcal{G}}^i_{[n/t_0]+1}.$$

Thus for any measurable function f that is constant on each $W^s \in \Gamma^s$, we have

$$\mathcal{F}^k \nu^1_k(f) = \mathcal{F}^k \nu^2_k(f).$$

This verifies the items (**C1**)(i)-(ii) in the Coupling Lemma 6.1

It follows from the estimation (6.15) that only the measure $\nu_1^i(\tau > n)$ dominates the amount of uncoupled measures at the $[n/t_0]$-th step. Thus using our new notations, we have proved item (**C1**)(iv):

$$\bar{\nu}_n^i(\mathcal{M}) = \nu^i(\tau > n + N_1) + \mathcal{O}(\nu^i(n + N_1 < \tau \leq n + N_1 + t_0)),$$

for any $n \geq 0$.

Next we prove (**C1**)(iii) by estimating $\mathcal{F}_*^n \bar{\nu}_n^1(f)$, for $f \in \mathcal{H}^-(\gamma_0)$. From now on, we choose the large constant $b = b(\gamma_0, \Lambda, \vartheta, \alpha_0) > 1$ such that

$$\Lambda^{-\gamma_0 b \ln n} < n^{-1-\alpha_0}, \quad \vartheta^{b \ln n} < n^{-1-\alpha_0}, \tag{6.17}$$

where $\vartheta \in (0, 1)$ is given by (4.5).

Next we estimate $\bar{\nu}_n^i$, which is enough to estimate $\nu^i(\tau > n)$. First we claim that

$$\bar{\nu}_n^i(C_{n,b}^c \cap (R \leq n)) \leq Cn^{-\alpha_0}. \tag{6.18}$$

To see this, note that points in $C_{n,b}^c \cap (R \leq n)$ will mostly visit cells with small indices and return to M at least ψ times within n iterations. We prove this claim by considering two cases.

(**a**). Let A_n be all points $x \in C_{n,b}^c \cap (R \leq n)$, such that the iterations of x hits $\mathcal{R}^*$ at most $b \ln n$ times within n iterations. Then there exists $k \in [1, n - b \ln n]$, such that $\mathcal{F}^k x \in M$ and the forward trajectory of $\mathcal{F}^k x$ return to $M \setminus \mathcal{R}^*$ at least $b \ln n$ consecutive times. Thus by the Coupling Lemma 4.6 for the induced map, $\mathcal{F}^k x$ be long to the base of the uncoupled measure for $F^{b \ln n}$. Since $\mathcal{G}^i$ are standard families and Lemma 4.6 implies that the coupling time Γ_M for the induced map $F\colon M \to M$ has exponential tail bound for standard families. It follows that, by (6.17),

$$\bar{\nu}_n^1(C_{n,b}^c \cap (R \leq n) \cap A_n) \leq n\|g^1\|_\infty \mu(x \in M \,:\, \Gamma_M > b \ln n) \leq Cn\|g^1\|_{C^\gamma} \vartheta^{b \ln n} \leq C\|g^1\|_{C^\gamma} n^{-\alpha} \tag{6.19}$$

(**b**). Now we consider points in $(C_{n,b}^c \cap (R \leq n)) \setminus A_n$. Then we claim that for any $x \in (C_{n,b}^c \cap (R \leq n)) \setminus A_n$, iterations of x hit $\mathcal{R}^*$ at least $b \ln n$ times within the $(b \ln n)^2$ returns to M. This is true because otherwise there must be an interval of length $b \ln n$ in these ψ returns to M such that iterates of x never hit $\mathcal{R}^*$, and this contradicts the assumption that (**a**) does not hold. More precisely, assume there exists $k \geq 1$, such that $\mathcal{F}^k x \in \mathcal{R}^*$, and there are at least $b \ln n$ more returns to $\mathcal{R}^*$ along the trajectory of $\{\mathcal{F}^m x, m = 1, \cdots, n\}$. This means that the forward images of the unstable manifold $W^u(x)$ have components properly returned to $\mathcal{R}^*$ at least $b \ln n$ times, and they all contain images of x. Then there exists $k \in [1, n - b \ln n]$, such that $\mathcal{F}^k x \in M$ and the forward trajectory of $\mathcal{F}^k x$ return to $M \setminus \mathcal{R}^*$ at least $b \ln n$ consecutive times. Thus by the Coupling Lemma 4.6 for the induced map, $\mathcal{F}^k x$ be long to the base of the uncoupled measure for $F^{b \ln n}$. Since $\mathcal{G}^i$ are standard families and Lemma 4.6 implies that the coupling time for the induced map $F\colon M \to M$ has exponential tail bound for standard families. It follows that, by (6.17),

$$\bar{\nu}_n^1(C_{n,b}^c \cap (R \leq n) \setminus A_n) \leq Cn\|g^1\|_{C^\gamma} \vartheta^{b \ln n} \leq C\|g^1\|_{C^\gamma} n^{-\alpha} \tag{6.20}$$

This finished the proof of our claim.

Thus the amount of uncoupled measure is essentially dominated by $C_{n,b} \cup (R > n)$, which contains points that returned to M fewer than $\psi = (b \ln n)^2$ times within n iterations. Thus by Assumption (**H2**), we know that

$$\mu(C_{n,b}) \leq Cn^{1-\alpha_0}.$$

Therefore,

$$\nu^i(C_{n,b}) \leq C\|g^i\|_\infty n^{1-\alpha_0},$$

where $d\nu^i = g^i d\mu$.

Combining the above facts, we have shown that for any $n \geq 1$,

$$\begin{aligned}
|\mathcal{F}_*^n \bar{\nu}_n^i(f)| &\leq C_1\|f\|_\infty n^{-\alpha_0} + |\bar{\nu}_n^i(f \circ \mathcal{F}^n \cdot \mathbf{I}_{C_{n,b}\cup(R>n)})| \\
&\leq C_1\|f\|_\infty n^{-\alpha_0} + \|f\|_\infty \|g^i\|_\infty \mu((C_{n,b} \cup (R > n))) \\
&\leq C\|g^i\|_\infty \|f\|_\infty n^{1-\alpha_0}.
\end{aligned}$$

This verifies **(C1)** (iii).

Note that if we take $\nu^1 = \mu$, then we observe that

$$\bar{\mu}_n(\mathcal{M}) \geq \mu(R > n) > cn^{1-\alpha_0}$$

for some constant $c > 0$. Combining the above estimates, we know that $\bar{\mu}_n(\mathcal{M}) \sim n^{1-\alpha_0}$, for any $n > \bar{N}_2$, for some large $\bar{N}_2 > N_1$. Since the set $(R > n)$ dominates the uncoupled measure, thus the coupled measure at the n-th iteration satisfies

$$\mathcal{F}_*^n \mu_n(\mathcal{R}^*) = \bar{\mu}_{n-1}(\mathcal{M}) - \bar{\mu}_n(\mathcal{M}) \leq \mu(M_{n-1}) + \mu(C_{n,b} \cap \mathcal{F}^{-n} M) \leq Cn^{1-\alpha_0+\varepsilon_1}.$$

By the mixing property, we know that $\mathcal{F}^n \nu^i \to \mu$ weakly, which implies that there exists $N_2 > \bar{N}_2$ large enough, such that for $n > N_2$,

$$\mathcal{F}_*^n \nu_n^i < Cn^{1-\alpha_0+\varepsilon_1},$$

which implies that **(C2)** also holds for n large enough.

7 Proof of Theorem 1.

Now it is time to investigate the rates of decay of correlations using the above Coupling lemma. We first prove a lemma that will be used later.

We consider any two standard families $\mathcal{G}^i = (\mathcal{W}^i, \nu^i)$ with probability density $g^i = d\nu^i/d\mu \in \mathcal{H}^+(\gamma_2)$ for $i = 1, 2$. And consider any $f \in \mathcal{H}^-(\gamma_1)$, with $\gamma_1, \gamma_2 \geq$

γ_0. According to the Coupling Lemma 6.1, there exist $N_1 = N_1(g^1, g^2) \geq 1$, such that we can define $\nu_0^i = \mathcal{F}_*^{N_1}\nu^i$; and for any $n \geq 1$, there exists a decompositions

$$\nu_0^i = \sum_{m=1}^{n} \nu_m^i + \bar{\nu}_n^i,$$

for $n > 1$, where $\mathcal{F}^m\nu_m^1$ is coupled with $\mathcal{F}^m\nu_m^2$. This implies that for any $m \geq 1$, for any measurable function $h \in \mathcal{H}^-(\gamma_1)$ that is constant on any stable manifold $W^s \in \Gamma^s$, we have

$$\mathcal{F}_*^m\nu_m^1(h) = \mathcal{F}_*^m\nu_m^2(h). \tag{7.1}$$

Lemma 7.1. *There exists $C_1 > 0$, which does not depend on f and g^i, $i = 1, 2$, such that for any $n \geq 1$,*

$$\left|\sum_{m=1}^{n}(\nu_m^i(f) - \nu_m^2(f))\right| \leq C_1\|f\|_{C^{\gamma_1}}^- \max\{\|g^1\|_\infty, \|g^2\|_\infty\}n^{1-\alpha_0-\varepsilon_1}.$$

Proof. For any $n \geq 1$, by definition of $C_{n,b}^c$ and the choice of b, for $f \in \mathcal{H}^-(\gamma_1)$ there exists $C > 0$ such that for $x \in C_{n,b}^c$ and $y \in W^s(x)$,

$$|f(\mathcal{F}^n(x)) - f(\mathcal{F}^n(y))| \leq C\|f\|_{\gamma_1}^-\Lambda^{-\gamma_1 b \ln n} \leq C\|f\|_{\gamma_1}^- n^{-\alpha_0}. \tag{7.2}$$

Now for any $x \in W^s \subset \Gamma^s$, we choose $\bar{x} \in W^s(x)$, such that $f(\bar{x}) = \max_{y\in W^s(x)} f(y)$ be the maximum value of f along stable manifold $W^s(x)$. Then (7.2) implies that for $x \in C_{n-m,b}^c \cap \Gamma^s$,

$$|f \circ \mathcal{F}^{n-m}(x) - f \circ \mathcal{F}^{n-m}(\bar{x})| \leq C\|f\|_{\gamma_1}^-(n-m)^{-\alpha_0}.$$

Thus we have for $i = 1, 2$,

$$\begin{aligned}
I_n^i :&= \sum_{m=1}^{n-1}\int_{\Gamma^s\cap C_{n-m,b}^c}\left|f \circ \mathcal{F}^{n-m}(x) - f \circ \mathcal{F}^{n-m}(\bar{x})\right| d\mathcal{F}_*^m\nu_m^i(x) \\
&\leq C\|f\|_{\gamma_1}^- \sum_{m=1}^{n-1}\mathcal{F}_*^m\nu_m^i(\Gamma^s)(n-m)^{-\alpha_0} \\
&\leq C\|f\|_{\gamma_1}^- \left(\sum_{m=1}^{n/2}\mathcal{F}_*^m\nu_m^i(\Gamma^s)(n-m)^{-\alpha_0} + \sum_{m=n/2}^{n-1}\mathcal{F}_*^m\nu_m^i(\mathcal{R}^*)(n-m)^{-\alpha_0}\right) \\
&\leq C_1\|f\|_{\gamma_1}^- \max\{\mathcal{F}_*^n\nu_n^i(\Gamma^s), n^{-\alpha_0}\} \leq C_1\|f\|_{\gamma_1}^-\|g^i\|_\infty n^{1-\alpha_0+\varepsilon_1},
\end{aligned}$$

where we have used **(C2)** in the Coupling Lemma 6.1 in the last estimate, as well as the following estimate:

$$\int_1^{n-1}\frac{1}{x^l}\cdot\frac{1}{(n-x)^t}\,dx \leq C_1 n^{-t} + C_2\frac{\ln n}{n^{l+t-1}} \leq Cn^{-t}, \tag{7.3}$$

for any $l \geq t \geq 1$.

Now we consider for $i = 1, 2$,

$$\begin{aligned}
I\!I_n^i :&= \sum_{m=1}^{n} \left| \int_{\Gamma^s \cap C_{n-m,b}} f \circ \mathcal{F}^{n-m}(x) d\mathcal{F}_*^m \nu_m^i(x) \right| \\
&\leq C\|f\|_\infty \sum_{m=1}^{n} \mathcal{F}_*^m \nu_m^i(C_{(n-m,b)}) \\
&\leq C\|f\|_\infty \sum_{m=1}^{n} (n-m)^{1-\alpha_0-\varepsilon_1} \mathcal{F}_*^m \nu_m^i(M) \\
&\leq C_1 \|f\|_\infty \|g^i\|_\infty n^{1-\alpha_0-\varepsilon_1},
\end{aligned}$$

where we have used Lemma 4.9.

Combining the above estimates, we have

$$I_n^1 + I_n^2 + I\!I_n^1 + I\!I_n^2 \leq C_1 \|f\|_{C^{\gamma_1}}^- \max\{\|g^1\|_\infty, \|g^2\|_\infty\} n^{1-\alpha_0-\varepsilon_1}.$$

This implies that

$$\begin{aligned}
&\sum_{m=1}^{n} \left| \int_{\mathcal{W}_m^1} f \circ \mathcal{F}^n d\nu_m^1 - \int_{\mathcal{W}_m^2} f \circ \mathcal{F}^n d\nu_m^2 \right| \\
&\leq \sum_{m=1}^{n} \left| \int_{\mathcal{F}^m \mathcal{W}_m^1} f \circ \mathcal{F}^{n-m} - f \circ \mathcal{F}^{n-m}(\bar{x}) d\mathcal{F}_*^m \nu_m^1 - \int_{\mathcal{F}^m \mathcal{W}_m^2} f \circ \mathcal{F}^{n-m} - f \circ \mathcal{F}^{n-m}(\bar{x}) d\mathcal{F}_*^m \nu_m^2 \right| \\
&\quad + \sum_{m=1}^{n} \left| \int_{\mathcal{F}^m \mathcal{W}_m^1} f \circ \mathcal{F}^{n-m}(\bar{x}) d\mathcal{F}_*^m \nu_m^1 - \int_{\mathcal{F}^m \mathcal{W}_m^2} f \circ \mathcal{F}^{n-m}(\bar{x}) d\mathcal{F}_*^m \nu_m^2 \right| \\
&\leq (I_n^1 + I_n^2 + I\!I_n^1 + I\!I_n^2) + \sum_{m=1}^{n} \left| \int_{\mathcal{F}^m \mathcal{W}_m^1} f \circ \mathcal{F}^{n-m}(\bar{x}) d\mathcal{F}_*^m \nu_m^1(x) - \int_{\mathcal{F}^m \mathcal{W}_m^2} f \circ \mathcal{F}^{n-m}(\bar{x}) d\mathcal{F}_*^m \nu_m^2(x) \right| \\
&= I_n^1 + I_n^2 + I\!I_n^1 + I\!I_n^2 \leq C_1 \|f\|_{C^{\gamma_1}}^- \max\{\|g^1\|_\infty, \|g^2\|_\infty\} n^{1-\alpha_0-\varepsilon_1}.
\end{aligned}$$

□

Using this lemma we can estimate the following decay rates of correlations.

Lemma 7.2. *There exist $C_1, C_2, C_3 > 0$, such that for any two standard families $\mathcal{G}^i = (\mathcal{W}^i, \nu^i)$ with $g^i = d\nu^i/d\mu \in \mathcal{H}^+(\gamma_2)$ for $i = 1, 2$, for any $f \in \mathcal{H}^-(\gamma_1)$, with $\gamma_1, \gamma_2 \geq \gamma_0$, $\nu^1(\mathcal{M}) = \nu^2(\mathcal{M})$, we have*

$$\left| \int_{\mathcal{M}} f \circ \mathcal{F}^n d\nu^1 - \int_{\mathcal{M}} f \circ \mathcal{F}^n d\nu^2 \right| \leq C\|f\|_{C^{\gamma_1}}^- \max\{\|g^1\|_{C^{\gamma_2}}^+, \|g^2\|_{C^{\gamma_2}}^+\}(\mu(C_{n,b}) + \mu(R > n)),$$

for any $n \geq N_1$, where $N_1 = N_1(g^1, g^2) \geq 1$.

Proof. By Lemma 6.1 there exist $N_1 = N_1(g^1, g^2) \geq 1$, for

$$\nu_0^i := \mathcal{F}_*^{N_1} \nu^i,$$

and for any $n \geq 1$, there exists a decomposition

$$\nu_0^i = \sum_{m=1}^{n} \nu_m^i + \bar{\nu}_n^i,$$

where $\mathcal{F}_*^m \nu_m^1$ is coupled with $\mathcal{F}^m \nu_m^2$ such that for any $m \geq 1$, for any measurable function $f \in \mathcal{H}^-(\gamma_1)$ that is constant on any stable manifold $W^s \in \Gamma^s$,

$$\mathcal{F}_*^m \nu_m^1(f) = \mathcal{F}_*^m \nu_m^2(f). \tag{7.4}$$

Using the above Lemma, we get

$$\begin{aligned}&\sum_{m=1}^{n} \left| \int_{\mathcal{W}_m^1} f \circ \mathcal{F}^n d\nu_m^1 - \int_{\mathcal{W}_m^2} f \circ \mathcal{F}^n d\nu_m^2 \right| \\ &\leq C_1 \|f\|_{C^{\gamma_1}}^- (\|g^1\|_\infty + \|g^2\|_\infty) n^{1-\alpha_0-\varepsilon_1}.\end{aligned}$$

Now, we denote $\nu_0^i := \mathcal{F}_*^{N_1} \nu^i$, then combining the above facts together with the Coupling Lemma 6.1, for both $i = 1, 2$,

$$\int_{\mathcal{M}} f \circ \mathcal{F}^n d\nu_0^1 - \int_{\mathcal{M}} f \circ \mathcal{F}^n \, d\nu_0^2 = \bar{\nu}_n^1(f \circ \mathcal{F}^n) - \bar{\nu}_n^2(f \circ \mathcal{F}^n) + \sum_{m=1}^{n} \left(\int_{\mathcal{W}_m^1} f \circ \mathcal{F}^n d\nu_m^1 - \int_{\mathcal{W}_m^2} f \circ \mathcal{F}^n d\nu_m^2 \right). \tag{7.5}$$

Our analysis implies that the second term is of order $n^{1-\alpha_0-\varepsilon_1}$, thus the decay rate is essentially dominated by $\bar{\nu}_n^1(f \circ \mathcal{F}^n)$ for general observable f. Note that

$$\begin{aligned}|\bar{\nu}_n^1(f \circ \mathcal{F}^n)| &= \bar{\nu}_n^1(f \circ \mathcal{F}^n \cdot \mathbf{I}_{C_{n,b} \cup (R>n)}) + \bar{\nu}_n^1(f \circ \mathcal{F}^n \cdot \mathbf{I}_{C_{n,b}^c \cap (R \leq n)})| \\ &\leq \|f\|_\infty \|g^1\|_\infty \left(\mu((C_{n,b} \cup (R > n))) + \mathcal{O}(n^{-\alpha_0}) \right). \end{aligned} \tag{7.6}$$

Now we can use (6.18) in the Coupling Lemma 6.1 for the next step estimation. Mainly because the uncoupled measure $\bar{\nu}_n^i(\mathcal{M})$ is dominated by $C_{n,b} \cup (R > n)$, while

$$\bar{\nu}_n^i(C_{n,b}^c) = \mathcal{O}(n^{-\alpha_0}).$$

Thus we have for $n \geq 1$,

$$\begin{aligned}&\left| \int_{\mathcal{M}} f \circ \mathcal{F}^{n+N_1} d\nu^1 - \int_{\mathcal{M}} f \circ \mathcal{F}^{n+N_1} \, d\nu^2 \right| \\ &\leq C_1 \|f\|_{C^{\gamma_1}}^- \max\{\|g^1\|_{C^{\gamma_2}}^+, \|g^2\|_{C^{\gamma_2}}^+\} (\mu(C_{n,b}) + \mu(R > n) + C_2 n^{1-\alpha_0-\varepsilon_1}) \\ &\leq 2C_1 \|f\|_{C^{\gamma_1}}^- \max\{\|g^1\|_{C^{\gamma_2}}^+, \|g^2\|_{C^{\gamma_2}}^+\} (\mu(C_{n,b}) + \mu(R > n)),\end{aligned}$$

where we have used Lemma 6.1 for the last step. This leads to the desired estimate as we have claimed. □

Next we consider the case when we do not have standard families, but only Hölder observables.

Lemma 7.3. *There exist $C > 0$ such that for any piecewise Hölder continuous functions $f \in \mathcal{H}^-(\gamma_1), g^i \in \mathcal{H}^+(\gamma_2)$, with $\gamma_i \geq \gamma_0$ and $\mu(g^1) = \mu(g^2)$, $i = 1, 2$, such that for any $n > 1$,*

$$|\mu(f \circ \mathcal{F}^{n+N_1} \cdot g^1) - \mu(f \circ \mathcal{F}^{n+N_1} \circ g^2)| \leq C\|f\|^-_{C^{\gamma_1}} \max\{\|g^1\|^+_{C^{\gamma_2}}, \|g^2\|^+_{C^{\gamma_2}}\} n^{1-\alpha_0},$$

where $N_1 = N_1(g^1, g^2) \geq 1$ is a large constant depending on g^1, g^2.

Proof. Without loss of generality, we can assume that

$$g^i = g_i^+ - g_i^-$$

is the decomposition of g^i into its positive and negative parts, with $\mu(g_i^\pm) > 0$. So both parts are piecewise dynamically Hölder continuous: $g_i^\pm \in \mathcal{H}^+(\gamma_2)$.

We start with the positive part g^+, which induces a probability measure

$$\nu_{i,+} = \frac{g_i^+}{\mu(g_i^+)}\mu$$

and gives rise to a standard family $(\mathcal{W}^u, \nu_+)$, for $i = 1, 2$. Thus by Lemma 7.2, there exists $N_1^+ := N_1^+(g_1^+, g_2^+)$, for any $f \in \mathcal{H}^-(\gamma_1)$, we have for any $n \geq 1$,

$$\left|\int_{\mathcal{M}} f \circ \mathcal{F}^{n+N_1^+} d\nu_{i,+} - \int_{\mathcal{M}} f \circ \mathcal{F}^{n+N_1^+} d\mu\right| \leq \frac{C\|f\|^-_{C^{\gamma_1}} \|g_i^+\|^+_{C^{\gamma_2}}}{\mu(g_i^+)} n^{1-\alpha_0}.$$

Similarly, there exists $N_1^- := N_1^-(g_1^-, g_2^-)$ such that for any $n > 1$, we get

$$\left|\int_{\mathcal{M}} f \circ \mathcal{F}^{n+N_1^-} d\nu_{i,-} - \int_{\mathcal{M}} f \circ \mathcal{F}^{n+N_1^-} d\mu\right| \leq \frac{C\|f\|^-_{C^{\gamma_1}} \|g_i^-\|^+_{C^{\gamma_2}}}{\mu(g_i^-)} n^{1-\alpha_0}.$$

Note that

$$g^i d\mu = \mu(g_i^+) d\nu_{i,+} - \mu(g_i^-) d\nu_{i,-}$$

Thus we actually get for all $n \geq 1$, with $N_1 := \max\{N_1^+, N_1^-\}$,

$$\left|\int_{\mathcal{M}} f \circ \mathcal{F}^{n+N_1} g^i d\mu - \mu(g^i)\mu(f)\right| \leq C_1\|f\|^-_{C^{\gamma_1}} \|g^i\|^+_{C^{\gamma_2}} n^{1-\alpha_0},$$

where $C_1 > 0$ is uniform constant. Thus if $\mu(g^1) = \mu(g^2)$, then we have

$$\begin{aligned}
&\left|\int_{\mathcal{M}} f \circ \mathcal{F}^{n+N_1} g^1 d\mu - \int_{\mathcal{M}} f \circ \mathcal{F}^{n+N_1} g^2 d\mu\right| \\
&\leq \left|\int_{\mathcal{M}} f \circ \mathcal{F}^{n+N_1} g^1 d\mu - \mu(g^1)\mu(f)\right| + \left|\int_{\mathcal{M}} f \circ \mathcal{F}^{n+N_1} g^2\, d\mu - \mu(g^2)\mu(f)\right| \\
&\leq C_1\|f\|^-_{C^{\gamma_1}} \max\{\|g^1\|^+_{C^{\gamma_2}}, \|g^2\|^+_{C^{\gamma_2}}\} n^{1-\alpha_0}.
\end{aligned}$$

□

Note that Theorem 1 directly follows from the above lemma.

8 Proof of Theorem 2.3 and Theorem 3.

8.1 Proof of Theorem 2

By (3.2), we denote

$$B_n = \bigcup_{m>n} \bigcup_{k=1}^{m-n} \mathcal{F}^k M_m = (R > n) \setminus M,$$

as the set of points in M^c that take at least n-iterations under $\mathcal{F}$ before they come back to M. We first prove the following lemma that will be used to prove Theorem 2.3.

Lemma 8.1. *For any large n, we define $\mu^n = \frac{\mu|_{B_n^c}}{\mu(B_n^c)}$. Then for any probability measure ν with support contained in M, we have for any bounded function f on M,*

$$\mathcal{F}_*^n \nu(f) - \mu(f) - \mu(f)\mu(R > n) = \mathcal{F}_*^n \nu(f) - \mathcal{F}_*^n \mu^n(f) + \mathcal{O}(n^{-\beta}), \tag{8.1}$$

with $\beta = \min\{\alpha_0, 2\alpha_0 - 2\}$.

Proof. First note that

$$B_n = \bigcup_{m>n} \bigcup_{k=1}^{m-n} \mathcal{F}^k M_m = (R > n) \setminus (\cup_{m \geq n} M_m),$$

which implies that

$$\mu(B_n) = \mu(R > n) + \mathcal{O}(n^{-\alpha_0}). \tag{8.2}$$

Since the support of the initial measure ν is contained in M, we have that after n iterations the push forward measure $\mathcal{F}_*^n \nu$ can never reach the region $\mathcal{F}^n B_n$. Thus we can ignore the measure μ restricted on B_n within first n-iterations. This fact implies that for each $n \geq 1$ the measure μ is a linear combination of two probability measures,

$$\mu = \mu(B_n) \cdot \frac{\mu|_{B_n}}{\mu(B_n)} + \mu(B_n^c) \cdot \frac{\mu|_{B_n^c}}{\mu(B_n^c)}.$$

We denote $\mu^n = \frac{\mu|_{B_n^c}}{\mu(B_n^c)}$. Note that for any bounded function f supported on M, we have

$$\mu(f) = \mu^n(f)\mu(B_n^c) = \mu^n(f) - \mu^n(f)\mu(B_n).$$

Notice also that (8.2) implies that $\mu(B_n) = O(n^{1-\alpha_0})$. Thus we have

$$\begin{aligned} \mu^n(f) &= \frac{\mu(f \cdot \mathbf{I}_{B_n^c})}{\mu(B_n^c)} = \frac{\mu(f)}{1 - \mu(B_n)} \\ &= \mu(f)(1 + \mu(B_n) + O(\mu(B_n)^2)). \end{aligned}$$

Thus

$$\mu^n(f)\mu(B_n) = \mu(f)(\mu(B_n) + O(\mu(B_n)^2)),$$

which also implies that

$$\mathcal{F}_*^n\mu^n(f)\mu(B_n) = \mathcal{F}_*^n\mu(f)(\mu(B_n) + O(\mu(B_n)^2)) = \mu(f)(\mu(B_n) + O(\mu(B_n)^2)).$$

Combining above facts, we have

$$\begin{aligned}\mathcal{F}_*^n\nu(f) - \mu(f) - \mu(f)\mu(R > n) &= \mathcal{F}_*^n\nu(f) - \mathcal{F}_*^n\mu^n(f) + O(\mu(R > n)^2) + \mathcal{O}(n^{-\alpha_0}) \\ &= \mathcal{F}_*^n\nu(f) - \mathcal{F}_*^n\mu^n(f) + \mathcal{O}(n^{-\beta}),\end{aligned}$$

with $\beta = \min\{\alpha_0, 2\alpha_0 - 2\}$, where we have used the fact given by (8.2).

□

To prove Theorem 2.3, we first assume the observable g defines a probability measure on M, with $d\nu = gd\mu$. By Lemma 8.1, we have

$$\mathcal{F}_*^n\nu(f) - \mu(f) - \mu(f)\mu(R > n) = \mathcal{F}_*^n\nu(f) - \mathcal{F}_*^n\mu^n(f) + O(n^{-\beta}). \tag{8.3}$$

Our goal is to show that for systems satisfy (**H**2) the correlation $\mathcal{F}_*^n\nu(f) - \mathcal{F}_*^n\mu^n(f) = o(\mu(R > n))$. We use assumption **(H2)**, which states that there exists $\varepsilon_1 > 0$ such that

$$\mu(C_{n,b} \cap \mathcal{F}^{-n}M) = \mathcal{O}(n^{1-\alpha_0-\varepsilon_1}).$$

Next we start to prove Theorem 2. Let $\gamma_1, \gamma_2 \geq \gamma_0$. We consider any $f \in \mathcal{H}^-(\gamma_1)$ supported on M, and any proper family $\mathcal{G} = (\mathcal{W}, \nu)$ with $g = d\nu/d\mu \in \mathcal{H}^+(\gamma_2)$ supported on M. For any large n, we define $\mathcal{G}^1 = \mathcal{G}$ and $\mathcal{G}^2 = (\mathcal{W}^u, \mu^n)$, then they are both proper families. We denote $\nu^1 = \nu$ and $\nu^2 = \mu^n$.

First note that since both measures are essentially supported on $(R \leq n)$, thus we have

$$\nu^i(R > n) = \mathcal{O}(n^{-\alpha}).$$

According to Lemma 7.1 we know that there exists $N = N(g)$, such that for any $k = 1, \cdots, n$,

$$\begin{aligned}&\left|\int_{\mathcal{M}} f \circ \mathcal{F}^k d\mathcal{F}_*^N\nu^1 - \int_{\mathcal{M}} f \circ \mathcal{F}^k d\mathcal{F}_*^N\nu^2\right| \\ &= \left|\bar{\nu}_k^1(f \circ \mathcal{F}^k) - \bar{\nu}_k^2(f \circ \mathcal{F}^k) + \sum_{m=1}^{k}\left(\int_{\mathcal{W}_m^1} f \circ \mathcal{F}^k d\nu_m^1 - \int_{\mathcal{W}_m^2} f \circ \mathcal{F}^k d\nu_m^2\right)\right| \\ &\leq \left|\bar{\nu}_k^1(f \circ \mathcal{F}^k) - \bar{\nu}_k^2(f \circ \mathcal{F}^k)\right| + C\|f\|_\infty\|g\|_\infty k^{1-\alpha_0+\varepsilon_1}.\end{aligned}$$

According to the Coupling Lemma 6.1 and (7.6), we know that the uncoupled measure $\bar{\nu}_k^i(\mathcal{M})$ is dominated by $C_{k,b}$, while $\bar{\nu}_k^i(C_{k,b}^c \cap (R \leq k)) = \mathcal{O}(k^{-\alpha_0})$. Thus

$$\begin{aligned}\bar{\nu}_k^i(f \circ \mathcal{F}^k) &= \bar{\nu}_k^i(f \circ \mathcal{F}^k \cdot \mathbf{I}_{C_{k,b}\cup(R>n)}) + \bar{\nu}_k^i(f \circ \mathcal{F}^k \cdot \mathbf{I}_{C_{k,b}^c}) = \bar{\nu}_k^1(f \circ \mathcal{F}^k \cdot \mathbf{I}_{C_{k,b}}) + \mathcal{O}(k^{-\alpha_0}) \\ &\leq \|f\|_\infty\|g^i\|_\infty\left(\mu(C_{k,b} \cap \mathcal{F}^{-k}\mathrm{supp}(f) \cap \mathrm{supp}(g^i))\right) + \mathcal{O}(k^{-\alpha_0}) \\ &= \|f\|_\infty\|g^i\|_\infty\left(\mu(C_{k,b} \cap \mathcal{F}^{-k}M \cap M)\right) + \mathcal{O}(k^{-\alpha_0}) = \mathcal{O}(k^{1-\alpha_0+\varepsilon_1}),\end{aligned}$$

where we have used Assumption (**H2**) in the last estimate. Combining above facts, we take $k = n$, then

$$|\int_{\mathcal{M}} f \circ \mathcal{F}^n d\mathcal{F}_*^N \nu^1 - \int_{\mathcal{M}} f \circ \mathcal{F}^n d\mathcal{F}_*^N \nu^2| \leq C_1 \|f\|_{C^{\gamma_1}}^- \|g\|_{C^{\gamma_2}}^+ n^{1-\alpha_0-\varepsilon_1}.$$

One can check that (**H1**) implies that $\mu(R > n + N_1 | R > n) = 1 + \mathcal{O}(n^{-1})$. Combining this with (8.3), then for any $n > N$, we get

$$|\mathcal{F}_*^n \nu(f) - \mu(f \circ \mathcal{F}^n) - \mu(f)\mu(B_n)| \leq |\mathcal{F}_*^n \nu(f) - \mu^n(f \circ \mathcal{F}^n)| + C_1 \|f\|_\infty \|g\|_\infty n^{-\beta}$$
$$\leq C\|f\|_\infty n^{1-\alpha_0-\varepsilon_1} + C_1 \|f\|_\infty \|g\|_\infty n^{-\beta},$$

here $\beta = \min\{2\alpha_0 - 2, \alpha_0\}$. For general Hölder observable g, we denote $g = g^+ - g^-$. It is enough to consider the case when $\mu(g^\pm) \neq 0$, then we consider g^+ and g^- separately as in the proof of Theorem 1, to get (2.14) in Theorem 2.

8.2 Prove of Theorem 3.

One interesting application of Theorem 2 is that one gets classical Central limiting theorem for stochastic processes generated by certain observables for dynamical systems with decay rates of order $\mathcal{O}(1/n)$.

We consider an observable $f \in H^-(\gamma_0) \cap H^+(\gamma_0)$ with $\text{supp}(f) \subset M$ supported in M. We assume f is not a coboundary. For any $n \geq 1$, we consider the two partial sums $\tilde{S}_n(f) := f + f \circ \mathcal{F} + \cdots + f \circ \mathcal{F}^n$ and $S_n(f) = f + f \circ F + \cdots + f \circ F^n$. We assume $\mu(f) = 0$. For the induced map (F, M, μ_M), it follows from Theorem 7.52 in [26] and [33], that conditions **(h1)-(h4)** implies that

$$\frac{S_n}{\sigma\sqrt{n}} \to N(0,1), \tag{8.4}$$

in distribution, where $N(0,1)$ is the standard normal variable, and by the Green-Kubo formula,

$$\sigma^2 = \mu_M(f^2) + 2\sum_{n=1}^{\infty} \mu_M(f \circ F^n \cdot f). \tag{8.5}$$

Note that

$$\int f \circ F^n \cdot f \, d\mu_M \leq C\|f\|_{C^{\gamma_0}} \vartheta^n$$

as the induced map enjoys exponential decay of correlations.

In [5], the partial sum S_n was associated with the so-called induced observable, $\tilde{S}_n f(x) = \sum_{m=0}^{R(x)-1} f(\mathcal{F}^m x)$. However, since our observable f has support contained in M, thus $\tilde{S}_n(f) = S_n(f)$ coincide with the induced observable. Next we review the relation between CLT of S_n and $\tilde{S}_n$, see a detailed proof in [5].

Lemma 8.2. *For any $n \geq 1$, if S_n satisfies the CLT (8.4), then $\tilde{S}_n$ also satisfy a CLT:*

$$\frac{\tilde{S}_n}{\tilde{\sigma}\sqrt{n}} \to N(0,1), \tag{8.6}$$

in distribution. Here $\tilde{\sigma}^2 = \sigma^2 \mu(M)$.

Thus (8.6) implies that $\tilde{S}_n$ satisfies the classical CLT with variance $\sigma\sqrt{\mu(M)}$. Moreover, again by the Green-Kubo formula we get

$$\sigma^2\mu(M) = \mu(f^2) + 2\sum_{n=1}^{\infty}\mu(f\circ\mathcal{F}^n\cdot f).$$

Comparing with (8.5), and use the definition $d\mu_M = d\mu/\mu(M)$, we get the interesting relation

$$\sum_{n=1}^{\infty}\mu(f\circ\mathcal{F}^n\cdot f) = \sum_{n=1}^{\infty}\mu(f\circ F^n\cdot f).$$

This finishes the proof of Theorem 3.

9 Sufficient conditions for (H2).

In this section, we introduce two sufficient conditions to guarantee assumption (**H2**).

Condition (H2)(a) . We assume there exist $\zeta\in(0,1), \eta_0\in(0,1)$, $C, C_1 > 0$, $N > 1$, such that for any sufficiently large $n > N$, and for each $m = 1,\ldots,(b\ln n)^2$,

$$\mu\big(\{x\in M\colon R(F^m(x)) > \varepsilon_n^m n\}|M_n\big) < C\eta_0^m n + C_1 n^{1-\zeta},$$

where $\varepsilon_n\in(0,1)$.

Condition (H2)(b). We assume there exist $q\in(0,1), C>0, p>0$, and $N>1$ such that for any sufficiently large $n>N$, and for each $m=1,\ldots,(b\ln n)^2$:

$$\mu\big(\{x\in M\colon R(F^m(x)) > n^{1-q}\}|M_n\big) < Cn^{-p}.$$

We will prove the following two lemmas.

Lemma 9.1. *Condition (**H**2)(a) implies that there exists $\varepsilon_1\in(0,1)$ such that **(H2)** holds.*

Proof. For any sufficiently large n, any $x\in C_{n,b}$, we define

$$k_0 = k_0(x) = \sum_{m=0}^{n-1}\mathbf{I}_M(\mathcal{F}^m(x))$$

as the number of returns to M within n iterations under $\mathcal{F}$ along the forward trajectory of x. By the definition of $C_{n,b}\subset(R<n)$, we know that any $x\in C_{n,b}$ must returns to M at least once, with

$$2\le k_0 < (b\ln n)^2$$

Let $x_n \in M$ be the last enter to M within n iterations of x, and we define

$$m_1(x) := \max_{1\le k\le k_0} \{R(F^{-k}(x_n))\}$$

to be the largest return time function value of R along n iterations of x under $\mathcal{F}$. i.e. there exists $k_1 = k_1(x) \in \{1, \cdots, n\}$, such that

$$x_1 := \mathcal{F}^{-k_1} x_n \in M_{m_1}.$$

Moreover we define indices:

$$m_1(x) = R(x_1), m_2(x) := R(F(x_1)), \cdots, m_{k_2}(x) := R(F^{k_2-1}(x_1)),$$

with $F^{k_2-1}(x_1) = F^{-1}x_n$ being the second last return to M along n iterations of x. Without loss of generality, we assume

$$\sum_{k=1}^{k_2} m_k \ge n/2,$$

i.e. the largest index m_1 occurs within the first $n/2$ iterations of x under $\mathcal{F}$. Then by assumption, since $m_1(x)$ is the largest index within k_0 returns to M along forward n iterations of x, we have

$$n/2 \le m_1 + \cdots + m_{k_2} \le n.$$

Since points $x \in C_{n,b}$ return to M at most ψ times during the first n iterates of $\mathcal{F}$, there are $\le \psi$ number of intervals between successive returns to M, and hence the longest interval has length $m_1(x) \ge n/\psi$. Let $c_0 < 1/2$ be a constant. We split $C_{n,b} = C'_{n,b} \cup C''_{n,b}$ into two disjoint parts.

(I) $C'_{n,b}$ is a 'good part' of $C_{n,b}$ such that for any $y \in C'_{n,b}$,

$$m_k(y) < \varepsilon_n^k m_1(y), \tag{9.1}$$

for all

$$\frac{c_0 n}{m_1(y)} \le k \le k_0 \le \psi(n) = (b \ln n)^2,$$

where $\varepsilon_n \in (0,1)$ was given in (**H2(a)**). More precisely, for $y \in C'_{n,b}$, there is a sequence of returns to M_{m_k}'s, with index decreasing exponentially in k, within n-iterations. For 'good' points $y \in C'_{n,b}$ we have

$$\frac{n}{2} \le \sum_{k=1}^{\psi(n)} m_k \le m_1 \cdot \frac{c_0 n}{m_1} + m_1 \sum_{k=\frac{c_0 n}{m_1}}^{\psi(n)} \varepsilon_n^k \le c_0 n + m_1 \varepsilon_n^{\frac{c_0 n}{m_1}} \le c_0 n + m_1.$$

Since $c_0 < 1/2$, we conclude $m_1 \ge cn$ for a positive constant $c := \frac{1}{2} - c_0 > 0$. This implies that for points in $C'_{n,b}$, the largest index m_1 within n iterations must be approximately of order n. Accordingly,

$$C'_{n,b} \cap \mathcal{F}^{-n} M \subset \bigcup_{m_1 \ge cn} \bigcup_{k=1}^{\psi(n)} \mathcal{F}^{-n}(F^k M_{m_1}).$$

Thus the measure of good points in $C_{n,b}$ is bounded by

$$\mu(C'_{n,b} \cap \mathcal{F}^{-n}M) \leq (b \ln n)^2 \sum_{m_1 \geq cn} \mu(M_{m_1}) \leq Cn^{-\alpha_0}(b \ln n)^2 \tag{9.2}$$

and

$$\mu(C'_{n,b}) \leq n \sum_{m_1 \geq cn} \mu(M_{m_1}) \leq Cn^{1-\alpha_0}.$$

(II) On the other hand $C''_{n,b}$ consists of 'bad' points $y \in C_{n,b}$, such that (9.1) fails. i.e.,

$$m_k(y) > \varepsilon_n^k m_1(y), \tag{9.3}$$

for some $\frac{c_0 n}{m_1(y)} \leq k \leq \psi(n)$.

We divide $C''_{n,b}$ according to the maximal index m_1 of its points

$$C''_{n,b} = \bigcup_{m_1 > \frac{n}{\psi}} C_{n,b,m_1},$$

such that C_{n,b,m_1} contains all points in $C''_{n,b}$ with the largest return time m_1 within n iterations. Note that

$$C_{n,b,m_1} \cap \mathcal{F}^{-n}M \subset \bigcup_{k=1}^{\psi(n)} \mathcal{F}^{-n}(F^k M_{m_1}).$$

The contribution of M_{m_1} to these 'bad' points in $C''_{n,b}$ will be estimated according to (**H**2)(**a**), which implies the following:

$$\mu(C_{n,b,m_1} \cap \mathcal{F}^{-n}M) \leq \psi(n)\mu(M_{m_1})(\eta_0^{c_0 n/m_1} + O(m_1^{-1+\zeta})),$$

where $\zeta > 0, \eta_0 \in (0,1)$ were defined in (**H**2).

By assumption $m_1 \geq n/\psi$, so the total measure of $C''_{n,b}$ is

$$\begin{aligned}\mu(C''_{n,b} \cap \mathcal{F}^{-n}M) &\leq \psi(n) \sum_{m_1 = n/\psi}^{n} \mu(M_{m_1})(\eta_0^{c_0 n/m_1} + O(m_1^{-1+\zeta})) \\ &\leq C_1(b \ln n)^{2+\alpha_0} n^{-\alpha_0} + C_2 \frac{\psi^{\alpha_0+2-\zeta}}{n^{\alpha_0+1-\zeta}} \leq \frac{C}{n^{\alpha_0+1-\varepsilon_1}},\end{aligned} \tag{9.4}$$

for some $\varepsilon_1 \in (0,1)$.

$$\mu(C''_{n,b}) \leq n \sum_{m_1 = n/\psi}^{n} \mu(M_{m_1})(\eta_0^{c_0 n/m_1} + O(m_1^{-1+\zeta})) \leq Cn^{1-\alpha_0}. \tag{9.5}$$

Combining the above estimates together with (9.2) we know that $C'_{n,b}$ dominates. Thus we get

$$\mu(C_{n,b} \cap \mathcal{F}^{-n}M) \leq Cn^{1-\alpha_0-\varepsilon_1}, \quad \mu(C_{n,b}) < Cn^{1-\alpha_0}.$$

□

Lemma 9.2. *Under assumption **(H2)(b)**, **(H2)** holds for any $\varepsilon_1 \in (0,1)$.*

Proof. For any sufficiently large n, any $x \in C_{n,b}$, we define

$$k_0 = k_0(x) = \sum_{m=0}^{n-1} \mathbf{I}_M(\mathcal{F}^m(x))$$

as the number of returns to M within n iterations under $\mathcal{F}$ along the forward trajectory of x. By the definition of $C_{n,b} \subset (R \le n)$, we know that any $x \in C_{n,b}$ must returns to M at least once, with $1 \le k_0 < (b \ln n)^2$. Let $x_n \in M$ be the last entrance to M within n-iterations, and we define

$$m_1(x) := \max_{1 \le k \le k_0 - 1} \{R(F^{-k}(x_n))\}$$

to be the largest return time function value of R along n iterations of x under $\mathcal{F}$. i.e. there exists $k_1 = k_1(x) \in \{1, \cdots, n\}$, such that $x_1 := \mathcal{F}^{-k_1} x_n \in M_{m_1}$, and $k_1 + m_1 \le n$. Moreover we define indices:

$$m_1(x) = R(x_1), m_2(x) := R(F(x_1)), \cdots, m_{k_2}(x) := R(F^{k_2 - 1}(x_1)),$$

with $F^{k_2-1}(x_1) = F^{-1}x_n$ being the second last return to M along n iterations of x. Without loss of generality, we assume

$$\sum_{k=1}^{k_2} m_k \ge n/2$$

i.e. the largest index m_1 occurs within the first $n/2$ iterations of x under $\mathcal{F}$. Then by assumption, since $m_1(x)$ is the largest index within k_0 returns to M along forward n iterations of x, we have

$$n/2 \le m_1 + \cdots + m_{k_2} \le n.$$

Since points $x \in C_{n,b}$ return to M at most ψ times during the first n iterates of $\mathcal{F}$, there are $\le \psi$ number of intervals between successive returns to M, and hence the longest interval has length $m_1(x) \ge n/\psi$. Let $c_0 < 1/2$ be a constant. We split $C_{n,b} = C'_{n,b} \cup C''_{n,b}$ into two disjoint parts.

(I) $C'_{n,b}$ is a 'good part' of $C_{n,b} \cap \mathcal{F}^{-n}M$ such that for any $y \in C'_{n,b}$,

$$m_k(y) < m_1(y)^{1-q}, \tag{9.6}$$

for all

$$\frac{c_0 n}{m_1(y)} \le k \le k_0 \le \psi(n) = (b \ln n)^2$$

More precisely, for $y \in C'_{n,b}$, there is a sequence of returns to M_{m_k}'s, with index decreasing exponentially in k, within n-iterations. For 'good' points $y \in C'_{n,b}$ we have

$$\frac{n}{2} \le \sum_{k=1}^{\psi(n)} m_k \le m_1 \cdot \frac{c_0 n}{m_1} + \sum_{k=\frac{c_0 n}{m_1}}^{\psi(n)} m_1^{1-q} \le c_0 n + m_1 \le c_0 n + m_1.$$

Since $c_0 < 1/2$, we conclude $m_1 \geq cn$ for a positive constant $c := \frac{1}{2} - c_0 > 0$. This implies that for points in $C'_{n,b}$, the largest index m_1 within n iterations must be approximately of order n. Accordingly,

$$C'_{n,b} \cap \mathcal{F}^{-n} M \subset \bigcup_{m_1 \geq cn} \bigcup_{k=1}^{\psi(n)} \mathcal{F}^{-n}(F^k M_{m_1}).$$

Thus the measure of good points in $C_{n,b}$ is bounded by

$$\mu(C'_{n,b} \cap \mathcal{F}^{-n} M) \leq (b \ln n)^2 \sum_{m_1 \geq cn} \mu(M_{m_1}) \leq C n^{-\alpha_0} (b \ln n)^2. \tag{9.7}$$

And

$$\mu(C'_{n,b}) \leq m \sum_{m_1 \geq cn} \mu(M_{m_1}) \leq C n^{1-\alpha_0}. \tag{9.8}$$

(II) On the other hand $C''_{n,b}$ consists of 'bad' points $y \in C_{n,b}$, such that (9.6) fails. i.e.,

$$m_k(y) > m_1(y)^{1-q}, \tag{9.9}$$

for some

$$\frac{c_0 n}{m_1(y)} \leq k \leq \psi(n).$$

We divide $C''_{n,b}$ according to the maximal index m_1 of its points

$$C''_{n,b} = \bigcup_{m_1 > \frac{n}{\psi}} C_{n,b,m_1},$$

such that C_{n,b,m_1} contains all points in $C''_{n,b}$ with the largest return time m_1 within n iterations. Note that

$$C_{n,b,m_1} \bigcap \mathcal{F}^{-n} M \subset \bigcup_{k=1}^{\psi(n)} \mathcal{F}^{-n}(F^k M_{m_1}).$$

The contribution of M_{m_1} to these 'bad' points in $C''_{n,b} \cap \mathcal{F}^{-n} M$ will be estimated according to (**H**2)(**b**), which implies the following:

$$\mu(C_{n,b,m_1} \cap \mathcal{F}^{-n} M) \leq \psi(n) \mu(M_{m_1}) m_1^{-p},$$

where $p > 0$ was defined in (**H**2).

By assumption $m_1 \geq n/\psi$, so the total measure of $C''_{n,b} \cap \mathcal{F}^{-n} M$ is

$$\mu(C''_{n,b} \cap \mathcal{F}^{-n} M) \leq \psi(n) \sum_{m_1 = n/\psi}^{n} \mu(M_{m_1}) m_1^{-p} \leq \frac{C}{n^{-\alpha_0}}. \tag{9.10}$$

Moreover,

$$\mu(C''_{n,b}) \leq n \sum_{m_1 = n/\psi}^{n} \mu(M_{m_1}) m_1^{-p} \leq \frac{C}{n^{1-\alpha_0}}. \tag{9.11}$$

Combining the above estimates together with (9.7) we know that $C'_{n,b}$ dominates. Thus we get

$$\mu(C_{n,b} \cap \mathcal{F}^{-n}M) \le C(b \ln n)^2 n^{-\alpha_0}, \quad \mu(C_{n,b}) \le C n^{1-\alpha_0}.$$

□

10 Applications to hyperbolic systems

To illustrate our method, we apply it to several classes of dynamical systems. Since the induced maps for most of these examples were studied in [32], we only remind some basic facts here.

First we recall standard definitions, see [12, 13, 17]. A 2-D flat billiard is a dynamical system where a point moves freely at unit speed in a domain $Q \subset \mathbb{R}^2$ and bounces off its boundary ∂Q by the laws of elastic reflection. We assume that $\partial Q = \cup_i \Gamma_i$ is a finite union of piecewise smooth curves, such that each smooth component $\Gamma_i \subset \partial Q$ is either convex inward (dispersing), or flat, or convex outward (focusing). Following Bunimovich, see [9, 11] and [26, Chapter 8], we assume that every focusing component Γ_i is an arc of a circle such that there are no points of ∂Q on that circle or inside it, other than the arc Γ_i itself. Under these assumptions the billiard dynamics is hyperbolic.

Let $\mathcal{M} = \partial Q \times [-\pi/2, \pi/2]$ be the *collision space*, which is a standard cross-section of the billiard flow. Canonical coordinates in $\mathcal{M}$ are r and φ, where r is the arc length parameter on ∂Q and $\varphi \in [-\pi/2, \pi/2]$ is the angle of reflection. The collision map $\mathcal{F}\colon \mathcal{M} \to \mathcal{M}$ takes an inward unit vector at ∂Q to the unit vector after the next collision, and preserves smooth measure $d\hat{\mu} = c \cdot \cos\varphi \, dr \, d\varphi$ on $\mathcal{M}$, where c is a normalization constant. Furthermore, $\partial\mathcal{M} \cup \mathcal{F}^{-1}(\partial\mathcal{M})$ is the singular set of $\mathcal{F}$.

For billiards with focusing boundary components, the expansion and contraction (per collision) may be weak during long series of successive reflections along certain trajectories. To study the mixing rates, one needs to find and remove the spots in the phase space where expansion (contraction) slows down. Such spots come in several types and are easy to identify, for example, see [28] and [26, Chapter 8]. Traditionally, we denote

$$\partial Q = \partial^0 Q \cup \partial^{\pm} Q,$$

where ∂Q^0 is the union of flat boundaries, ∂Q^- contains focusing boundaries and ∂Q^+ contains dispersing boundaries. The collision space can be naturally divided into focusing, dispersing and neutral parts:

$$\mathcal{M}_0 = \{(r, \varphi)\colon r \in \partial^0 Q\}, \qquad \mathcal{M}_{\pm} = \{(r, \varphi)\colon r \in \partial^{\pm} Q\}.$$

Now we define the induced phase space:

$$M = \{x \in \mathcal{M}_-\colon \ \pi(x) \in \Gamma_i, \, \pi(\mathcal{F}^{-1}x) \in \Gamma_j, \, j \ne i\} \cup \mathcal{M}_+. \tag{10.1}$$

Note that M only contains all collisions on dispersing boundaries and the *first* collisions with each focusing arc (the collisions with straight lines are skipped altogether). The map F preserves the measure μ conditioned on M, which we denote by $\mu = [\hat{\mu}(M)]^{-1}\hat{\mu}$.

Furthermore, F has a larger singular set than the original map. Let $S_0 = \partial M$, then $S_1 := S_0 \cup F^{-1}S_0$ is the singular set of F. Let $R\colon \mathcal{M} \to \mathbb{N}$ be the first return time function, such that for any $x \in \mathcal{M}$, $\mathcal{F}^{R(x)}x$ returns to M for the first time. We define $M_m = R^{-1}\{m\} \cap M$ as the level set of the return time function.

To be more specific, we consider billiard systems that have been studied in [28, 29, 30, 31, 27], which include semi-dispersing billiards on a rectangle, Bunimovich billiard, Bunimovich Stadia, skewed stadia, billiards with flat points, billiards with cusps. It was proved in these references that the induced system (F, M, μ) satisfies the condition **(H1)** and **(H3)** and enjoys exponential decay of correlations. It is enough to check (**H**2). We first introduce some new conditions that imply **(H2)** and which are easier to check.

10.1 Billiards with property (H2)(a).

New condition (A1). Assume for any n large, there exist $\hat{M}_n \subset M_n$, $\zeta_1 \in (0,1)$, such that

$$\mu(M_n \setminus \hat{M}_n | M_n) < C/n^{\zeta_1},$$

and for any $x \in \hat{M}_n$, there exist $d \in (0,1)$, $C > 0$ and a large $b > 0$, such that for $m = 1, \cdots, (b \ln n)^2$, we have the following condition:

$$\mathbb{E}\left[R(F^m(x)) | x \in \hat{M}_n\right] \leq d^m n + \mathcal{O}(n^{-1-\zeta_1}).$$

Proposition 10.1. Condition (**A1**) implies **(H2)(a)**.

Proof. Let $\psi(n) = (b \ln n)^2$. For any n large, any $k = 1, \cdots, \psi(n)$, consider the quotient

$$\xi_{n,k}(x) \colon = R(F^k x)/R(F^{k-1}x),$$

for any $x \in \hat{M}_n$. Let

$$S_{n,k} = \ln \xi_{n,k}(x) + \cdots + \ln \xi_{n,1}(x).$$

Then the moment generating function of S_k at 1 satisfies:

$$\mathbb{M}_k(1) := \mu_n(e^{S_{n,k}}) = \mu_n(\prod_{i=1}^{k} \xi_{n,i}) \leq d^k + \mathcal{O}(n^{-\zeta_1}),$$

where we have used assumption (**A1**) in the last step, and $\mu_n := \mu|_{M_n}/\mu(M_n)$. Now by the Markov inequality, for any $\rho > 0$

$$\mu_n(e^{S_{n,k}} < e^{-\rho k}) \leq e^{\rho k}\mu_n(e^{S_{n,k}}) \leq (de^{\rho})^k + \mathcal{O}(n^{-\zeta_1})e^{\rho k}.$$

We choose $q \in (0,1)$ such that $\zeta_1 + q < 1$, and define

$$\rho_n = \frac{q}{b^2 \ln n},$$

then one can check that there exists $N > 1$, such that

$$\eta_0 := \sup_{n \geq N} de^{\rho_n} < 1,$$

and $e^{\rho_n k} < n^q$, for any $k = 1, \cdots, \psi(n)$. Now we have

$$\mu_n(S_{n,k} > -\rho_n k) \leq \eta_0^k + \mathcal{O}(n^{-\zeta_1 + q}).$$

Let $\varepsilon_n = e^{-\rho_n}$, then above inequality is equivalent to

$$\mu_n(R(F^k(x)) > \varepsilon_n^k n) < \eta_0^k + \mathcal{O}(\frac{1}{n^{1-\delta}}).$$

Thus **(H2)(a)** holds, with $n\varepsilon_n^k \geq n^{1-q}$, for $k = 1, \cdots, (b \ln n)^2$, and $\delta := \zeta_1 - q \in (0,1)$. □

The stadium billiard table, introduced by Bunimovich in [9], is comprised of two equal semicircles which are connected by two parallel lines. Dynamics on the stadium have been shown to be non-uniformly hyperbolic, ergodic, and mixing; for some discussion of these facts see [9, 11, 27]. In [55] Markarian proved that correlations in Stadia decay as $\mathcal{O}((\ln n)^2/n)$. Chernov and Zhang later improved in [31] the decay rate to $\mathcal{O}(1/n)$.

It was shown in [11, 26] that if $x \in M_m$, then $Fx \in M_k$ for $k \in B_m = [a_m, b_m]$ with $a_m \asymp m/\beta$, $b_m \asymp \beta m$, where $\beta = 3$. Here $a_m \asymp \sqrt{m}$ means that there exist $c_1 > c_2 > 0$, such that $c_2 \leq \frac{a_m}{\sqrt{m}} \leq c_1$, for any $m \geq 1$. The transition probability between cells is given by

$$\mu_M(Fx \in M_k | x \in M_m) = \frac{3m}{8k^2} + \mathcal{O}\left(\frac{1}{m^2}\right).$$

It was also shown in [26, 28, 32] that for any $x \in M_n$,

$$\bar{\eta} := \mathbb{E}(\ln(\tfrac{R(F)}{R}) | R(x) = m) = \sum_{k \in B_m} \ln \frac{k}{m} \cdot \frac{3m}{8k^2} < 1 - \frac{5}{4} \ln 3 < 0.$$

In [31], condition (3.5)-(3.8) were checked for both Bunimovich Stadia and Skewed Stadia. Moreover, the remarks before Lemma 3.3 in [31] shows that there exists $\hat{M}_n \subset M_n$, with

$$\mu(M_n \setminus \hat{M}_n | M_n) \leq Cn^{-1/2}$$

and for any $x \in \hat{M}_n$,

$$\ln(R(F^m)/R(x)) \leq \eta_1 + \cdots + \eta_m,$$

where $\{\eta_1, \cdots, \eta_m\}$ are independent random variables with the same distribution as η. Here η is a random variable supported on $[-\ln\beta, \ln\beta+1]$, and having a negative mean value $\mathbb{E}(\eta) = \bar{\eta} < 0$. We define $d := e^{\bar{\eta}} < 1$.

Thus one can check that for any $x \in M_n$, for $t = 1, \cdots, (b\ln n)^2$,

$$\mathbb{E}(\tfrac{R(F^t)}{R}|R(x) = n) \leq \mathbb{E}(e^{\eta_1+\cdots+\eta_t}|R(x) = n) + \mathcal{O}(n^{-\frac{1}{2}}) \leq d^t + \mathcal{O}(n^{-\frac{1}{2}}).$$

This implies condition (**A1**), with $\zeta_1 = 1/2$. Now by Proposition 10.1, we know that condition (**A1**) implies **(H2)(a)**.

The case for skewed stadia is very similar to above analysis, so it satisfies **(H2)(a)**, which we will omit here.

10.2 Billiards with property (H2)(b).

Assume that each cell M_n has dimension $\asymp n^{-a}$ in the unstable direction, dimension $\asymp n^{-\beta}$ in the stable direction, and measure $\mu(M_n) \asymp n^{-d}$, with $d \geq a + \beta > 2$. We first foliate M_n with unstable curves $W_\alpha \subset M_n$ (where α runs through an index set $\mathcal{A}$). These curves have length $|W_\alpha| \asymp n^{-\beta}$. Let $\nu_n := \frac{1}{\mu(M_n)}\mu|_{M_n}$ be the conditional measure of μ restricted on M_n. Let $\mathcal{W} = \cup_{\alpha\in\mathcal{A}} W_\alpha$ be the collection of all unstable curves, which foliate the cell M_n. Then we can disintegrate the measure ν along the leaves W_α. More precisely, in this way we can obtain a standard family $\mathcal{G}_n = (\mathcal{W}, \nu_n)$, such that for any measurable set $A \subset M_n$,

$$\nu_n(A) = \int_{\mathcal{A}} \nu_\alpha(W_\alpha \cap A)\, d\lambda(\alpha),$$

where (W_α, ν_α) is a standard pair, and λ is the probability factor measure on $\mathcal{A}$. For some $k \leq n$, let $\mathcal{R}_k = \bigcup_{i>k} M_i$, which contains all the cells with index greater than k. For each unstable curve $W_\alpha \in \mathcal{W}$, if $F^n W_\alpha$ crosses $\mathcal{R}_k$, then $F^m W_\alpha$ is cut into pieces by the boundary of cells in $\mathcal{R}_k$. Moreover, the largest length of these pieces is $\sim k^{-\beta}$. According to the growth lemma (3.18), there exists $\theta_0 \in (0, 1)$, such that we have

$$F_*^m \nu_n(\mathcal{R}_k) \leq c\vartheta_0^m F_* \nu_n(\mathcal{R}_k) + C_z k^{-\beta q_0}. \tag{10.2}$$

Case I. Billiards with cusps.

This class of billiards were first studied by Machta [54]. It is known that the billiard maps on these tables are hyperbolic and ergodic. However, the hyperbolicity is non-uniform. As a result, correlations decay with order $\mathcal{O}(n^{-1})$, see [28, 31, 27]. Moreover, it was showed that it satisfies the One-Step Expansion **(h4)** with $q_0 = 1$.

In [27] Chernov and Markarian showed that the induced map F on a subset $M \subset \mathcal{M}$ has exponential decay of correlations. Dynamics of F on billiards with cusps are remarkably different than those on a stadium when it comes to points

travelling between m-cells: if $x \in M_m$ and $Fx \in M_k$, then $k \in B_m = [a_m, b_m]$, with $a_m \asymp \sqrt{m}, b_m \asymp m^2$. And the transition probability from the m-cell to the k-cell is

$$\mu_M(Fx \in M_k | x \in M_m) := \frac{\mu(\{x \in M_m : Fx \in M_k\})}{\mu(M_m)} \asymp \frac{m^{2/3}}{k^{7/3}},$$

with $k \in [\sqrt{m}, m^2]$. Each cell M_m has length approximately $m^{-7/3}$ in the unstable direction and length approximately $m^{-2/3}$ in the stable direction. Its measure is $\mu(M_m) \sim m^{-3}$.

Moreover, it was checked in [31] at the end of section 5 that this class of billiards satisfies for any small enough $e \in (0, 1/4)$:

$$\mu(R(F(x)) > n^{\frac{1}{2}+e} | R(x) = n) \leq Cn^{-\frac{1}{2e}},$$

for some uniform constant $C > 0$. Since each cell M_n has length approximately $n^{-7/3}$ in the unstable direction, we take $\beta = 7/3$, and $k = n^{\frac{1}{2}+e}$. Then we have

$$F_* \nu_n(\mathcal{R}_{n^{\frac{1}{2}+e}}) \leq Cn^{-\frac{1}{2e}}.$$

Now we apply (10.2) to get for any $i = 1, \cdots, (b \ln n)^2$,

$$\begin{aligned} \nu_n(R(F^i(x)) > n^{\frac{1}{2}+e}) &= F_*^i \nu_n(\mathcal{R}_k) \\ &\leq c\vartheta_0^i F_* \nu_n(\mathcal{R}_k) + C_z k^{-\beta} \\ &\leq C\vartheta_0^i n^{-\frac{1}{2e}} + C_z n^{-\frac{7}{3}(\frac{1}{2}+e)}. \end{aligned}$$

This verifies **(H2)**(b) with $q = 1/2 - e$, $p = \frac{1}{2e}$.

Case II. Semi-dispersing billiards. Billiards in a square with a finite number of fixed, disjoint circular obstacles removed are known as semi-dispersing billiards. Chernov and Zhang proved [31] that this system has a decay of correlations bounded by const $\cdot\, n^{-1}$. Here the reduced phase space M is made up only of collisions with the circular obstacles. The induced map $F : M \to M$ is then equivalent to the well studied Lorentz gas billiard map without horizon [28], which is known to have exponential decay of correlations (see [27], for instance). The structure of the m-cells $M_m = \{x \in M : R(x) = m\}$ is examined thoroughly in [12, 13, 27]. We will use some of the facts presented in those references. Many properties of the m-cells and of the induced billiard map in the semi-dispersing case are quite similar to those in billiards with cusps. In particular, the measure of each m-cell is again $\mu_M(M_m) \asymp m^{-3}$, with u-dimension approximately m^{-2}. Thus we take $\beta = 2$. Moreover it satisfies the One-Step Expansion Estimate **(h4)** with $q_0 = 1$. It is also know that for a point $x \in M_m$, $Fx \in M_k$ where $k \in B_m = [a_m, b_m]$, with

$$a_m \asymp \sqrt{m}, b_m \asymp m^2,$$

as in billiards with cusps. One major change is the transition probabilities between cells. For semi-dispersing billiards, we have (for admissible k) that

$$\mu_M(Fx \in M_k | x \in M_m) \asymp \frac{m+k}{k^3}.$$

Moreover, it was checked in [31] at Section 5 that this class of billiards satisfies for any small enough $e \in (0, 1/4)$:

$$\mu(R(F(x)) > n^{\frac{1}{2}+e} | R(x) = n) \leq Cn^{-\frac{1}{2e}},$$

for some uniform constant $C > 0$. We take $k = n^{\frac{1}{2}+e}$. Then we have

$$F_* \nu_n(\mathcal{R}_{n^{\frac{1}{2}+e}}) \leq Cn^{-\frac{1}{2e}}.$$

Now we apply (10.2) to get for any $i = 1, \cdots, (b \ln n)^2$,

$$\begin{aligned} \nu_n(R(F^i(x)) > n^{\frac{1}{2}+e}) &= F_*^i \nu_n(\mathcal{R}_k) \\ &\leq cF_* \nu_n(\mathcal{R}_k) + C_z k^{-\beta} \\ &\leq Cn^{-\frac{1}{2e}} + C_z n^{-2(\frac{1}{2}+e)}. \end{aligned}$$

This verifies **(H2)**(b) with $q = 1/2 - e$, $p = \frac{1}{2e}$.

This implies that the semi-dispersing billiards on a rectangle and dispersing billiards with cusps have optimal bounds of decay rates of correlations given by Theorem 2.3.

10.3 Application to linked-twist maps

In this section, we apply our main theorem to the linked-twist map studied in [67]. We claim that this map satisfy the new condition (**A1**)

We consider the two-dimensional torus $\mathbb{T}^2 = [0, 2) \times [0, 2)$ with coordinates (x, y) (mod 2). On this torus we define two overlapping annuli P, Q by $P = [0, 2] \times [0, 1]$, $Q = [0, 1] \times [0, 2]$. We denote the union of the annuli by $R = P \cup Q$ and the intersection by $M = P \cap Q$. The annuli P and Q are vertical and horizontal strips in the torus. In order to define a linked twist map on the torus we first define a twist map on each annulus. A twist map is simply a map in which the orbits move along parallel lines, but with a uniform shear. In particular, we define $F : R \to R$, such that

$$F(x, y) = \begin{cases} (x + 2y, y), & \text{if } (x, y) \in P; \\ (x, y), & \text{if } (x, y) \in R \setminus P. \end{cases}$$

Note that F leaves points in $R \setminus P$ unchanged, and any horizontal line in P is invariant. We define the map G similarly:

$$G(x, y) = \begin{cases} (x, y + 2x), & \text{if } (x, y) \in Q; \\ (x, y), & \text{if } (x, y) \in R \setminus Q. \end{cases}$$

Now the linked twist map H is defined by $H := G \circ F$, which maps from R to R. By calculating the differential DH, one can easily check that $\det DH = 1$, which implies that H preserves the Lebesgure measure m on M.

We will first define a reduced map which enjoys the exponential decay of correlations. More precisely, we define $F_M : M \to M$, to be the return map with respect to F, such that for any $(x, y) \in M$, $F_M(x, y) = F^n(x, y)$, where $n = R_F(x, y)$ is the first return time of (x, y) to M under iterations of F. Similarly, we define $G_M : M \to M$, such that $G_M(x, y) = G^n(x, y)$, where $n = R_G(x, y)$ is the first return of (x, y) to M under iterations of G. We define the reduced map as $T := G_M \circ T$. Then T is the first return map obtained from H onto M. Note that as G is an Anosov diffeomorphism restricted on M, so by the uniformly hyperbolicity of G on M, there exists $N = N(G) > 1$ such that $G^N M \subset M$. Let m_M be the conditional Lebesgure measure on M, then T preserves m_M.

Let $S_{\pm 1}$ be the singular set of the reduced map $H^{\pm 1} := H_S^{\pm 1}$. In [67], Figure 2 shows the structure of S_1 while Figure 5 shows the image of S_{-1}. Using the notation of that paper, we label by $\{\Sigma_n\}$ the connected regions near $(1, 0)$ in S_1, as shown in Figure 6, on which the return time function is n. We know from Appendix A of [67] that the cell Σ_n has length of order $1/n$ and width of order $1/n^2$. Similarly, we denote by $\{M_n\}$ the level set (or called cells) in S_{-1} with backward return time n. As it was shown in Lemma 5.3 of [67], unstable manifolds have slope $1 + \sqrt{2}$, thus we know that the longer boundary curves of M_n all have slope approximately $1 + \sqrt{2}$, and these cells converge to the fixed point $(1, 0)$ as $n \to \infty$. In addition, one can show that M_n has length of order $O(n^{-1})$ and width of order $O(n^{-2})$. In the proof of Lemma 5.4 of [67], it was shown that when an unstable manifold W intersects Σ_n for some n large enough, it only crosses those Σ_m with $m \in [n, (3+2\sqrt{2})n]$. If we redefine n, then we can say that W intersects only cells Σ_m, with $m \in I_n = [n/\beta + c_1, \beta n + c_1]$, where $\beta = 1 + \sqrt{2}$, for some constants c_1, c_2.

In terms of the singular set S_{-1}, this implies that the image of $\partial M_n \subset S_{-1}$ will only intersect Σ_m, for $m \in I_n$, i.e. $M_n \subset \cup_{m \in I_n} \Sigma_m$. Thus we take an unstable manifold W that completely stretches across M_n, then its image HW will be cut into pieces such that each piece is stretched completely across M_m, for $m \in I_n$.

Note that for large n, the region $M_n \cap \Sigma_n$ is nearly a rectangle with dimension $O(m^{-2}) \times O(n^{-2})$. Now Lemma 5.2 in [67] implies that the expansion factor of unstable manifolds in Σ_m is $O(m)$, thus $TM_n \cap \Sigma_m$ is a strip in M_m that completely stretched in the unstable direction and has width $O(\frac{1}{mn^2})$.

Thus one can now check that the transition probability of moving from Σ_n to Σ_m is

$$\frac{\mu(\Sigma_m \cap TM_n)}{\mu(TM_n)} = \frac{c_0 \frac{1}{m^2 n^2}}{\frac{1}{n^3}} = c_0 \frac{n}{m^2},$$

where $c_0 = \beta - \beta^{-1}$ is the normalizing constant, such that

$$\sum_{m \in I_n} \mu(\Sigma_m \cap TM_n) = \mu(TM_n)$$

More precisely, c_0 solves

$$\sum_{m=n/\beta}^{\beta n} c_0 \frac{1}{m^2 n^2} = \frac{1}{n^3}.$$

Thus we have shown that this class of maps satisfy (**A1**). By Proposition 10.1, the map satisfies condition (**H2**(a)). Thus Theorem 1-3 hold for this map.

11 Appendix. List of major notations

Here we list notations in the paper that have been used at least twice.

- $(\mathcal{F}, \mathcal{M}, \mu)$ — the original nonuniform hyperbolic system;
- (F, M, μ_M) — the induced uniformly hyperbolic system;
- $\mathcal{W}^u$ (resp. $\mathcal{W}^s$) — the collection of all unstable (resp. stable) manifolds for $\mathcal{F}$;
- $\mathcal{W}^u_F$ (resp. $\mathcal{W}^s_F$) — the collection of all unstable (resp. stable) manifolds for F;
- $R : \mathcal{M} \to \mathbb{N}$ — the first hitting time to M. It is an extension of the first return time function on M;
- M_m — the m-th level set of R in M, i.e. $M_m = \{x \in M : R(x) = m\}$, it is the closure of the open set $\mathcal{D}_m$, for $m \geq 1$;
- $\mathcal{S}_1$ (resp. $\mathcal{S}_{-1}$) — the singular set of $\mathcal{F}$ (resp. $\mathcal{F}^{-1}$);
- S_1 (resp. S_{-1}) — the singular set of F (resp. F^{-1});
- ρ_W — the density function of the u-SRB measure μ_W;
- (W, ν) — a standard pair;
- $(\mathcal{W}, \nu)$ — a standard family, also denoted as $\mathcal{G} = (\mathcal{W}, \nu)$, equipped with a factor measure λ on the index set $\mathcal{A}$ of $\mathcal{W}$;
- $F^n(\mathcal{W}, \nu)$ — a standard family defined as $(F^n\mathcal{W}, F^n_*\nu)$ with factor measure λ_n on the index set $\mathcal{A}^n$;
- $\{(W_\alpha, \mu_\alpha), \alpha \in \mathcal{A}^u_M, \lambda^u_M\}$ — the standard family $(\mathcal{W}^u_F, \mu_M)$;
- $\{(W_\alpha, \mu_\alpha), \alpha \in \mathcal{A}^u_{\mathcal{M}}, \lambda^u_{\mathcal{M}}\}$ — the standard family $(\mathcal{W}^u, \mu)$;
- $\mathfrak{F}(\mathcal{M})$ — the collection of all standard families on $\mathcal{M}$;
- $C_{n,b}$ — the set of all points in $\mathcal{M}$ whose forward orbits have returned to M at least once, but at most $(b \ln n)^2$ times within n iterations, defined as in (2.2);

- $\gamma_0 \in (0,1)$ – is the constant defined in the distortion bound (2.5);
- ε_1 — a constant defined in assumption (**H2**);
- $\mathcal{R}^* = \Gamma^u \cap \Gamma^s$ — the hyperbolic product set defined as in Definition 4.1;
- $\tilde{F}$ — higher iterations of F, i.e. $\tilde{F} = F^{n_1}$, as defined in (4.4);
- $\tilde{\mathcal{F}}$ — higher iterations of $\mathcal{F}$, i.e. $\tilde{\mathcal{F}} = \mathcal{F}^{n_2}$, as defined in Proposition 5.1;
- N_1 — the number of iterations needed for a standard family to become proper, defined as in lemma 5.2, N_1 depends on the choice of standard family;
- $T : \mathcal{M} \to \mathcal{R}^*$ — the first hitting time to $\mathcal{R}^*$, as defined in (5.1) such that $T = \mathcal{F}^\tau$;
- $\mathcal{D}_{n,m}$ — an s-subset in γ^s, such that $T^n \mathcal{D}_{n,m} = \mathcal{F}^m \mathcal{D}_{n,m}$, as defined in (5.3);
- $\alpha_0 > 1$ — the constant in (**H1**), defining the order of $\mu(R > n) \leq Cn^{1-\alpha_0}$;
- $\gamma_i \geq \gamma_0$, $i = 1, 2$ — the Hölder exponent of functions in $\mathcal{H}^{\pm}(\gamma_i)$;
- $C_{\mathbf{r}} > 0$ — the distortion constant defined in (**h3**);
- $C_F > C_{\mathbf{r}}$ — the constant in the definition of standard pair (3.4);
- $C_q > 100 C_F$ — the constant defining a proper family $\mathcal{G}$, i.e. $\mathcal{Z}(\mathcal{G}) < C_q$;
- δ_0 — a constant defined in (3.13), such that a standard pair (W, ν) is called a proper standard pair, if $|W| < 20\delta_0$;
- $\hat{\delta}_1 \in (0,1)$ — constants defined in Proposition 4.2;
- $\hat{\delta}_0 \in (0, \hat{\delta}_1)$ — a constant defined in Lemma 4.5;
- $\delta > 0$ — a constant defined in Lemma 5.2, and also used in Lemma 5.7;
- $\Lambda > 1$ — the minimal expansion factor given in assumption (**h1**);
- $t_0 \geq 1$ — the number of iterations for each step of coupling process chosen in Lemma 5.7;
- $\vartheta_0 \in (0,1)$ — a constant defined in the absolute continuity property (**h3**);
- $\vartheta \in (0,1)$ — a constant chosen in the coupling lemma for the induced map, i.e. Lemma 4.6;
- $\vartheta_1 = \max\{\vartheta_0, \Lambda^{-\gamma_0}\}$ — a constant in $(0,1)$ as chosen in Lemma 5.6;
- $\vartheta_2 = \max\{\vartheta, \lambda^{-\gamma_1}\}$ — a constant for the exponential decay rates of $\{f \circ F^n\}$ as in (4.7);
- $\vartheta_3 \in (0,1)$ — the constant defined in the Growth Lemma 3.5.

References

[1] J. Alves, Ch. Bonatti, M. Viana, *SRB measures for partially hyperbolic systems whose central direction is mostly expanding*, Inventiones Math. **140**, (2000).

[2] J. Alves, V. Pinheiro, *Slow rates of mixing for dynamical systems with hyperbolic structures*, J. Stat. Phys., **131** (2008), 505–534.

[3] V. Baladi, S. Gouezel, *Good Banach spaces for piecewise hyperbolic maps via interpolation*, Ann. Inst. H. Poincaré-Analyse non linéaire **26** (2009), 1453–1481.

[4] V. Baladi, C. Liverani, *Exponential decay of correlations for piecewise cone hyperbolic contact flows*, Comm. Math. Phys., **314** (2012), 689–773.

[5] P. Balint, N. Chernov and D. Dolgopyat, *Limit theorems for dispersing billiards with cusps*, Comm. Math. Phys., **308** (2011), 479–510.

[6] P. Bálint and S. Gouezel, *Limit theorems in the stadia billiards*, Comm. Math. Phys., **263** (2006), 451–512.

[7] P. Bálint and S. Gouezel, *Limit theorems in the stadia billiards*, Comm. Math. Phys., **263** (2006), 451–512.

[8] L. Barreira, Y. Pesin, *Nonuniformly Hyperbolicity: Dynamics of Systems with Nonzero Lyapunov Exponents*, Encyclopedia of Mathematics and Its Applications, Cambridge University Press, 2007.

[9] L. A. Bunimovich, *On billiards close to dispersing*, Math. USSR. Sb. **23** (1974), 45–67.

[10] C. Bose, R. Murray, *Polynomial decay of correlations in the generalized baker's transformation*, J. Bifur. Chaos (2013).

[11] L. A. Bunimovich, *On ergodic properties of nowhere dispersing billiards*, Comm. Math. Phys. **65** (1979), 295–312.

[12] L. A. Bunimovich, Ya. G. Sinai & N. I. Chernov, *Markov partitions for two-dimensional billiards*, Russ. Math. Surv. **45** (1990), 105–152.

[13] L. A. Bunimovich, Ya. G. Sinai & N. I. Chernov, *Statistical properties of two-dimensional hyperbolic billiards*, Russ. Math. Surv. **46** (1991), 47–106.

[14] L.A. Bunimovich& H. Spohn, *Viscosity for a periodic two disk fluid: an existence proof*, Comm. Math. Phys. **176** (1996), 661–680.

[15] X. Bressaud and C. Liverani *Anosov diffeomorphism and coupling*, Ergod. Th. Dynam. Syst., **22** (2002), 129–152.

[16] N. Chernov, *Limit theorems and Markov approximations for chaotic dynamical systems*, Prob. Th. Rel. Fields **101** (1995), 321–362.

[17] N. Chernov, *Decay of correlations and dispersing billiards*, J. Stat. Phys. **94** (1999), 513–556.

[18] N. Chernov, *Sinai billiards under small external forces*, Ann. H. Poincaré **2** (2001), 197–236.

[19] N. Chernov, *Sinai billiards under small external forces II*, Ann. H. Poincaré **9** (2008), 91–107.

[20] N. Chernov and D. Dolgopyat, *Brownian Brownian Motion – I*, Memoirs AMS, **198**, No 927 (2009), 193 pp.

[21] N. Chernov and D. Dolgopyat, *Galton Board: limit theorems and recurrence*, J. Amer. Math. Soc., **22** (2009), 821–858.

[22] N. Chernov and D. Dolgopyat, *Anomalous current in periodic Lorentz gases with infinite horizon*, Russ. Math. Surv. **64** (2009), 651–699.

[23] N. I. Chernov, G. L. Eyink, J. L. Lebowitz and Ya. G. Sinai, *Steady-state electrical conduction in the periodic Lorentz gas*, Comm. Math. Phys. **154** (1993), 569–601.

[24] N. Chernov and A. Korepanov, *Spatial structure of Sinai-Ruelle-Bowen measures*, Physica D, to appear.

[25] N. I. Chernov and J. L. Lebowitz, *Stationary shear flow in boundary driven Hamiltonian systems*, Phys. Rev. Lett. **75** (1995), 2831–2833.

[26] N. Chernov and R. Markarian, *Chaotic billiards*, Math. Surv. Monographs, **127**, AMS, Providence, RI (2006), 316 pp.

[27] N. Chernov and R. Markarian, *Dispersing billiards with cusps: slow decay of correlations*, Commun. Math. Phys., **270** (2007), 727–758.

[28] N. Chernov and H.-K. Zhang, *Billiards with polynomial mixing rates*, Nonlineartity, **4** (2005), 1527–1553.

[29] N. Chernov and H.-K. Zhang, *A family of chaotic billiards with variable mixing rates*, Stochast. Dynam. **5** (2005), 535–553.

[30] N. Chernov and H.-K. Zhang, *Regularity of Bunimovich's stadia*, Reg. Chaotic Dyn. **3** (2007), 335–356.

[31] N. Chernov and H.-K. Zhang, *Improved estimates for correlations in billiards*, Comm. Math. Phys., **277**, 2008, 305–321.

[32] N. Chernov and H.-K. Zhang, *On statistical properties of hyperbolic systems with singularities*, J. Stat. Phys., **136** (2009), 615–642.

[33] N. Chernov, H.-K. Zhang and P. Zhang *Electrical current in Sinai billiards under general small forces*, Journal of Statistical Physics, **153** (2013), 1065-1083.

[34] M. Denker *The central limit theorem for dynamical systems*, In: Dynamical Systems and Ergodic Theory, Banach Center Publ. **23**, Warsaw, 1989.

[35] M. Demers, C. Liverani, *Stability of Statistical Properties in Two-dimensional Piecewise Hyperbolic Maps*, TAMS **360** (2008), 4777–4814.

[36] M. Demers and H.-K. Zhang, *Spectral analysis of the transfer operator for the Lorentz Gas*, J. Modern Dyn. **5**, (2011), 665–709.

[37] M. Demers and H.-K. Zhang, *A functional analytic approach to perturbations of the Lorentz gas*, Comm. Math. Phys. **324** (2013), 767–830.

[38] M. Demers and H.-K. Zhang, *Spectral analysis of hyperbolic systems with singularities*, Nonlinearity, **27** (2014), 379–433.

[39] D. Dolgopyat, On dynamics of mostly contracting diffeomorphisms, *Comm. Math. Phys.* **213** (2001), 181–201.

[40] D. Dolgopyat, *Limit Theorems for partially hyperbolic systems*, TAMS, **356** (2004), 1637–1689.

[41] D. Dolgopyat, *On differentiability of SRB states for partially hyperbolic systems*, Invent. Math. **155** (2004), 389–449.

[42] B. Friedman, R. F. Martin, *Behavior of the velocity autocorrelation function for the periodic Lorentz gas*, Physica D **30** (1988), 219–227.

[43] A.O.Gelfond, *An estimate for the reminder term in a limit theorem for recurrent events*, Theory Prob. Applications, 9 (1964), 299-303.

[44] S. Gouezel, *Sharp polynomial estimates for the decay of correlations*, Israel J. Math., **139** (2004), 29–65.

[45] B. Hasselblatt and A. Katok, *Handbook of Dynamical Systems*, Volume 1B, Ed. by B. Hasselblatt and A. Katok, Elsevier, Amsterdam, 2005.

[46] H. Hu, *Decay of correlations for piecewise smooth maps with indifferent fixed points*, Ergod. Th. Dynam. Syst., **24** (2004), 495–524.

[47] H. Hu and S. Vaienti, *Absolutely Continuous Invariant Measures for Nonuniformly Expanding Maps*, Ergod. Th. Dynam. Syst., **29** (2009), 1185–1215.

[48] H. Hu and S. Vaienti, *Polynomial Decay for Non-uniformly Expanding Maps*, preprint, arXiv:1307.0359

[49] A.Katok, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Inst. Hautes tudes Sci. Publ. Math. 51 (1980), 137173.

[50] A. Katok, J.-M. Strelcyn, *Invariant Manifolds, Entropy and Billiards; Smooth Maps with Singularities.* Lect. Notes Math., **1222**, Springer, New York (1986) (with the collaboration of F. Ledrappier & F. Przytycki).

[51] A.N. Kolmogorov, *Markov chains with a countable number of possible states*, Byull.Mosk.Un-ta,(A), (1937) 1-16.

[52] C. Liverani, M. Martens, *Convergence to equilibrium for intermittent symplectic maps*, Communications in Mathematical Physics, **260** (2005), 527–556.

[53] C. Liverani, D. Terhesiu, *Mixing for some non-uniformly hyperbolic systems*, To appear in Annales Henri Poincare.

[54] J. Machta, *Power law decay of correlations in a billiard problem*, J. Statist. Phys. **32** (1983), 555–564.

[55] R. Markarian, *Billiards with polynomial decay of correlations*, Ergod. Th. Dyn. Syst. **24** (2004), 177–197.

[56] I. Melbourne, D. Terhesiu *Decay of correlations for nonuniformly expanding systems with general return times.* Ergodic Theory Dyn. Syst **34** (2014) 893-918

[57] Ya. Pesin, *Dynamical systems with generalized hyperbolic attractors: hyperbolic, ergodic and topological properties.* Ergod. Th. Dyn. Syst. **12**, 123–152 (1992).

[58] Ya. Pesin, Ya.G. Sinai *Hyperbolicity and stochasticity of dynamical systems.* Sov. Sci. Rev. Sect. C Math. Phys. Rev. **4**, 53–115 (1981).

[59] Y. Pomeau and P. Manneville, *Intermittent transition to turbulence in dissipative dynamical systems*, Comm. Math. Phys. **74** (1980), 189–197.

[60] B.A. Rogozin, *An estimate of the remainder term in limit theorems of renewal theorey*, Theory of probability and its applications, **4** (1973), 662-677.

[61] E. Sataev, *Invariant measures for hyperbolic maps with singularities.* Russ. Math. Surv. **47** (1992), 191–251.

[62] Ya. Pesin, Ya.G. Sinai, *Gibbs measures for partially hyperbolic attractors*, Ergod. Th. Dyn. Sys. **2** (1982), 417–438.

[63] O. Sarig, *Subexponential decay of correlations*, Invent. Math., **150** (2002), 629–653.

[64] Ya.G. Sinai, *Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards*, Uspehi Mat. Nauk **25** (1970), 141–192.

[65] C.Stone, *On characteristic functions and renewal theory*, Thans.Amer.Math.Soc., 120 (1965), 327-342.

[66] D. Szasz and T. Varju, *Limit Laws and Recurrence for the Planar Lorentz Process with Infinite Horizon*, J. Stat. Phys. **129** (2007), 59–80.

[67] J. Springham, R. Sturman, *Polynomial decay of correlations in linked-twist maps*, Ergod. Th. Dynam. Sys., to appear.

[68] S. Vaienti and H.-K. Zhang, *Optimal decay rates of correlations for general systems with hyperbolic invariant manifolds.* Preprint.

[69] F. Vivaldi, G. Casati and I. Guarneri, I., *Origin of long-time tails in strongly chaotic systems*, Phys. Rev. Lett. **51** (1983), 727–730.

[70] F. Wang, H.-K. Zhang, P-F. Zhang, *Decay of correlations for unbounded observables*, submitted.

[71] L.-S. Young, *Statistical properties of systems with some hyperbolicity including certain billiards*, Ann. Math. **147** (1998), 585–650.

[72] L.-S. Young, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188.

[73] L.-S. Young, *What are SRB measures, and which dynamical systems have them?*, J. Statist. Phys. **108** (2002), 733–751.