



Large deviation principle in one-dimensional dynamics

Yong Moo Chung¹ · Juan Rivera-Letelier² ·
Hiroki Takahasi³

Received: 23 July 2018 / Accepted: 9 June 2019 / Published online: 17 July 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract We study the dynamics of smooth interval maps with non-flat critical points. For every such a map that is topologically exact, we establish the full (level-2) Large Deviation Principle for empirical means. In particular, the Large Deviation Principle holds for every non-renormalizable quadratic map. This includes the maps without physical measure found by Hofbauer and Keller, and challenges the widely-shared view of the Large Deviation Principle as a refinement of laws of large numbers.

Mathematics Subject Classification 37A50 · 37C40 · 37D25 · 37D45 · 37E05

✉ Yong Moo Chung
chung@amath.hiroshima-u.ac.jp

Juan Rivera-Letelier
riveraletelier@gmail.com
http://rivera-letelier.org

Hiroki Takahasi
hiroki@math.keio.ac.jp
http://www.math.keio.ac.jp/~hiroki/

¹ Department of Applied Mathematics, Hiroshima University, Higashi-Hiroshima 739-8527, Japan

² Department of Mathematics, University of Rochester, Hylan Building, Rochester, NY 14627, USA

³ Department of Mathematics, Keio Institute of Pure and Applied Sciences (KiPAS), Keio University, Yokohama 223-8522, Japan

1 Introduction

An important concept in dynamical systems is that of *physical measure*. An invariant probability measure μ of a dynamical system f is *physical* if there exists a set E of positive Lebesgue measure in the phase space such that for every $x \in E$ the empirical mean on the orbit $\{x, f(x), f^2(x), \dots, f^{n-1}(x)\}$ converges to μ as $n \rightarrow \infty$, in the weak* topology. The theory of large deviations aims to provide exponential bounds on the probability that the empirical means stay away from μ . See, e.g., [14, 18] for general accounts of large deviation theory.

For uniformly hyperbolic diffeomorphisms, physical measures have been constructed in the pioneering works of Sinai, Ruelle and Bowen [4, 43, 47]. In this setting, the *Large Deviation Principle* (LDP for short) has been established by Takahashi [48, 49], Orey & Pelikan [35], Kifer [28], Young [50]; it describes stochastic features of deterministic dynamics with chaotic behavior.

In recent years there have been considerable efforts to extend these results beyond the uniformly hyperbolic setting. All previous results we are aware of are restricted to maps satisfying a weak form of hyperbolicity, see for example [8, 9, 12, 20, 26, 30, 33, 37, 41] and references therein. The only ones establishing a full LDPs are [8] and [9, Theorem B], for a set of positive measure of quadratic maps satisfying the Collet-Eckmann condition [11]. See also [12, 20, 30] for full LDPs for maps satisfying a weak form of hyperbolicity, in which the empirical measures are weighted with respect to an equilibrium state of a Hölder continuous potential. In spite of the relative incompleteness of the theory, there was a belief among experts that the LDP holds under weaker assumptions.

In this paper we study smooth interval maps with only non-flat critical points. The presence of critical points is a severe obstruction to uniform hyperbolicity. We establish a full level-2 LDP for every such map that is topologically exact. In particular, the LDP holds for every non-renormalizable quadratic map. Notably, this includes maps having no physical measure, like the quadratic maps found by Hofbauer & Keller in [21, 22]. Notice that the formulation of the LDP [17] does not *a priori* assume the strong law of large numbers or the existence of a physical measure.

We now proceed to describe our main results in more detail.

1.1 Statement of results

Throughout this paper we set $X = [0, 1]$, and for a measurable subset A of X we denote by $|A|$ its Lebesgue measure.

A *critical point* of a differentiable map $f: X \rightarrow X$ is a point at which the derivative of f vanishes. Denote by $\text{Crit}(f)$ the set of critical points of f .

A critical point c of f is *non-flat* if there are $\ell > 1$ and diffeomorphisms ϕ and ψ of \mathbb{R} such that $\phi(c) = \psi(f(c)) = 0$ and such that for every x in a neighborhood of c ,

$$|\psi \circ f(x)| = |\phi(x)|^\ell.$$

Note that a continuously differentiable map with only non-flat critical points has at most a finite number of critical points.

Denote by \mathcal{M} the space of Borel probability measures on X endowed with the weak* topology. For $x \in X$ denote by $\delta_x \in \mathcal{M}$ the Dirac measure at x . Given a continuous map $f: X \rightarrow X$ and an integer $n \geq 1$, define $\delta_x^n = \frac{1}{n} \sum_{i=0}^{n-1} \delta_{f^i(x)}$. The map f is *topologically exact* if for every nonempty open subset U of X there is an integer $n \geq 1$ such that $f^n(U) = X$.

Main Theorem *Let $f: X \rightarrow X$ have Hölder continuous derivative and only non-flat critical points. If f is topologically exact, then the full level-2 Large Deviation Principle holds, namely, there exists a lower semi-continuous function $I: \mathcal{M} \rightarrow [0, \infty]$ such that:*

– (lower bound) for every open subset \mathcal{G} of \mathcal{M} ,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log |\{x \in X : \delta_x^n \in \mathcal{G}\}| \geq -\inf_{\mathcal{G}} I;$$

– (upper bound) for every closed subset \mathcal{K} of \mathcal{M} ,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\{x \in X : \delta_x^n \in \mathcal{K}\}| \leq -\inf_{\mathcal{K}} I.$$

In the theorem above and in the rest of the paper,

$$\log 0 = -\infty, \inf \emptyset = \infty \text{ and } \sup \emptyset = -\infty.$$

The function I is called a *rate function*. From the general theory of large deviations [14, 18], the LDP determines I uniquely. We show that $-I$ is the upper semi-continuous regularization of the “free energy function”. Then the rate function is convex, and it is characterized as the Legendre transform of the cumulant generating function, see Sect. 1.2.

The traditional application of the LDP in dynamical systems is for maps having a physical measure. In the probabilistic viewpoint of dynamical systems, the existence of a physical measure is analogous to the law of large numbers, and the LDP is a refinement of this law. For concreteness, consider a map $f: X \rightarrow X$ as in the Main Theorem that in addition has a physical measure μ . Then the rate function I vanishes at μ and, assuming f is sufficiently regular, for Lebesgue almost every point x in X the sequence of empirical measures $\{\delta_x^n\}_{n=1}^\infty$ converges to μ in the weak* topology, see [6, Theorem 8]. This

last property is thus analogous to the law of large numbers, and the LDP given by the Main Theorem is a refinement: the speed of convergence is controlled by the rate function I .

The LDP given by the Main Theorem applies to situations beyond the traditional one, since it does not require the existence of a physical measure. Note also that the LDP in the Main Theorem does not require any weak form of hyperbolicity. To illustrate the broader applicability of the the Main Theorem, we give two new insights into the dynamics of quadratic maps. The first concerns one of the quadratic maps f_0 without physical measures studied by Hofbauer & Keller in [21, 22]. The rate function of f_0 vanishes entirely on its effective domain, in sharp contrast with the uniformly hyperbolic case where the rate function only vanishes at the physical measure. The LDP given by the Main Theorem gives a quantitative version of the “maximal oscillation” property studied by Hofbauer & Keller in [22], see Sect. 1.2 for details. We also consider the quadratic Fibonacci map f_* studied by Lyubich & Milnor [32], Keller & Nowicki [27], and others. The equilibrium states of f_* for the geometric potential $-\log |Df_*|$ form a segment, having the physical measure μ_* of f_* as an endpoint. Although the basin of an equilibrium state μ different from μ_* has zero Lebesgue measure, the LDP given by the Main Theorem implies that μ still attracts a significant set of initial conditions, see Sect. 1.2 for details.

Besides the uniformly hyperbolic case mentioned at the beginning of the introduction, the only previous full LDPs were established in [8] and [9, Theorem B] for a set of positive measure of quadratic maps satisfying the Collet-Eckmann condition. See also [12, 20, 30]¹ for full LDPs for maps satisfying a weak form of hyperbolicity, in which the empirical measures are weighted with respect to an equilibrium state of a Hölder continuous potential. For local LDPs, see [26, Theorems 1.2 and 1.3], [33], [37, Corollary B.4], [41], and references therein.

We now state a corollary of the Main Theorem that follows from the general theory of large deviations. We use it below to compare our result with previous related ones. Let $\mathcal{M}(f)$ be the subspace of \mathcal{M} of those measures that are f -invariant. For a continuous function $\varphi: X \rightarrow \mathbb{R}$ define

$$c_\varphi = \min \left\{ \int \varphi d\nu : \nu \in \mathcal{M}(f) \right\} \text{ and } d_\varphi = \max \left\{ \int \varphi d\nu : \nu \in \mathcal{M}(f) \right\},$$

and for each integer $n \geq 1$ and x in X write

$$S_n \varphi(x) = \sum_{i=0}^{n-1} \varphi \circ f^i(x) = n \int \varphi d\delta_x^n.$$

¹ See also the survey article of Denker [15].

Moreover, define a rate function $q_\varphi: \mathbb{R} \mapsto [0, \infty]$ by

$$q_\varphi(t) = \inf \left\{ I(\mu) : \mu \in \mathcal{M}, \int \varphi d\mu = t \right\}.$$

This function is bounded on $[c_\varphi, d_\varphi]$ and constant equal to ∞ on $\mathbb{R} \setminus [c_\varphi, d_\varphi]$. Furthermore, q_φ is convex on \mathbb{R} , and therefore continuous on (c_φ, d_φ) .

The following corollary is a direct consequence of the Main Theorem and of the contraction principle, see for example [14, 18].

Corollary *Let $f: X \rightarrow X$ have Hölder continuous derivative and only non-flat critical points. If f is topologically exact, then for every continuous function $\varphi: X \rightarrow \mathbb{R}$ satisfying $c_\varphi < d_\varphi$ and for every interval J intersecting (c_φ, d_φ) ,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \left| \left\{ x \in X : \frac{1}{n} S_n \varphi(x) \in J \right\} \right| = - \inf_J q_\varphi.$$

One previous result relevant to this corollary is that of Keller & Nowicki [26, Theorem 1.2], in the case where f is a S -unimodal map satisfying the Collet-Eckmann condition, see the definition of S -unimodal map below. Denoting by μ_{ac} the unique absolutely continuous invariant probability (*acip* for short) of f , they proved that the corollary holds with $\varphi = \log |Df|$ for every interval J whose boundary is contained in a small neighborhood of $t = \int \log |Df| d\mu_{\text{ac}}$.

Let us illustrate a broad applicability of the Main Theorem and its corollary in the context of “ S -unimodal” maps, which we proceed to recall. A non-injective continuously differentiable map $f: X \rightarrow X$ is *unimodal*, if $f(\partial X) \subset \partial X$, and if f has a unique critical point. The unique critical point c of such a map must be in the interior of X and be of “turning” type; that is, f is not locally injective at c . The map f is *S -unimodal*, if in addition c is non-flat for f , and if f is of class C^3 and has negative Schwarzian derivative on $X \setminus \{c\}$; in this context the non-flatness condition is the same as above with the additional requirement that the diffeomorphisms ϕ and ψ are of class C^3 .

Each S -unimodal map has exactly one of the following dynamical characteristics:

- (i) it has an attracting cycle;
- (ii) it is infinitely renormalizable;
- (iii) it is at most finitely renormalizable.

In case (iii) there is an integer $p \geq 1$ and a closed interval J containing the critical point of f in its interior, such that $f^p(J) \subset J$, such that the return map $f^p: J \rightarrow J$ is topologically exact, and such that the intervals $J, f(J), \dots, f^{p-1}(J)$ have mutually disjoint interiors, see for example the combination of [13, Theorem V.1.3] and [45, Theorem 2.19 and Proposition 2.34].

This implies that a rescaling of $f^p|_J$ satisfies the assumptions of the Main Theorem. It follows that *the LDP holds for every at most finitely renormalizable S -unimodal map*.

For a real analytic family of S -unimodal maps with quadratic critical point and non-constant combinatorics, such as the quadratic family, Lebesgue almost every parameter corresponds to either case (i) or case (iii), and in the latter case there is an acip [1, 31]. The set of parameters corresponding to acips has positive Lebesgue measure [2, 23].

1.2 Further results and comments

We characterize the rate function I in the Main Theorem as follows. For $\nu \in \mathcal{M}(f)$ denote by $h(\nu)$ the entropy of ν , and define the *Lyapunov exponent* $\lambda(\nu)$ of ν by $\lambda(\nu) = \int \log |Df| d\nu$. The *free energy function* $F: \mathcal{M} \rightarrow [-\infty, \infty)$ is defined by,

$$F(\nu) = \begin{cases} h(\nu) - \lambda(\nu) & \text{if } \nu \in \mathcal{M}(f); \\ -\infty & \text{otherwise.} \end{cases}$$

Since the map f in the Main Theorem is topologically exact, it has the specification property. Then it has no hyperbolic attracting periodic point and empirical measures along periodic orbits are dense in the space of invariant measures [46, Theorem 1]. Together with the upper semi-continuity of the Lyapunov exponent, this implies that for every $\nu \in \mathcal{M}(f)$ we have $\lambda(\nu) \geq 0$, see also [42, Proposition A.1]. We show that the rate function I in the Main Theorem is given by

$$I(\mu) = - \inf_{\mathcal{G} \ni \mu} \sup_{\mathcal{G}} F, \quad (1)$$

where the infimum is taken over all open subsets \mathcal{G} of \mathcal{M} containing μ . It follows that I is convex, and therefore that I is the Legendre transform of the cumulant generating function, see for example [14, Theorem 4.5.10(b)]. On the other hand, using (1) and the fact that the rate function takes only nonnegative values, we obtain from the LDP in the Main Theorem that for every $\nu \in \mathcal{M}(f)$ we have $F(\nu) \leq 0$. This is known as Ruelle's inequality [44]. Note also that the rate function vanishes at each equilibrium state of f for the geometric potential $-\log |Df|$. That is, the rate function I vanishes at every measure $\nu \in \mathcal{M}(f)$ for which Rohlin's formula $F(\nu) = 0$ holds. See below for an example where the function vanishes at a measure that is not an equilibrium state.

Consider a S -unimodal map f with a non-flat critical point that satisfies the Collet-Eckmann condition [11]. Then the corresponding rate function vanishes precisely at the (unique) acip [10, Theorem A.1]. As mentioned earlier, for such

a map f we have the traditional application of the LDP in the Main Theorem as a refinement of the law of large numbers.

We now describe two applications of the LDP in the Main Theorem that go beyond the traditional application of refining the law of large numbers. First, we consider one of the quadratic maps $f_0: X \rightarrow X$ without physical measures studied by Hofbauer & Keller in [21, Theorem 5] and [22], see Theorem A.1 in the Appendix for a precise description. The Main Theorem applies to f_0 and the corresponding rate function vanishes entirely on its effective domain, see Theorem A.2 in the Appendix. This is in sharp contrast with the uniformly hyperbolic case, for which the rate function only vanishes at the physical measure. Applying the Corollary of the Main Theorem to f_0 , we obtain:

Choose $\varepsilon > 0$, an arbitrary invariant measure μ , and an arbitrary continuous function $\varphi: X \rightarrow \mathbb{R}$. Then for $n \geq 1$, the set E_n of all the initial conditions x_0 for which

$$\left| \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f_0^j(x_0)) - \int \varphi \, d\mu \right| \leq \varepsilon,$$

is sub-exponentially large with respect to n :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log |E_n| = 0. \quad (2)$$

Equivalently, there is a sub-exponentially large set of initial conditions for which the Birkhoff average of φ is near the mean with respect to μ . This happens simultaneously for every invariant measure μ , and gives a quantitative version of the “maximal oscillation” property of f_0 shown by Hofbauer & Keller in [22].

The second application is for the Fibonacci quadratic map $f_*: X \rightarrow X$, studied by Lyubich & Milnor [32], Keller & Nowicki [27], and others. This map has a physical measure μ_* whose basin of attraction has full Lebesgue measure on X [32, Theorem 1.3(4)]. That is, for Lebesgue almost every point x in X the sequence of empirical measures $\{\delta_x^n\}_{n=1}^\infty$ converges to μ in the weak* topology. On the other hand, the closure of the critical orbit is a Cantor set that supports a unique invariant probability measure ν_* [32, Theorem 1.2]. The measures μ_* and ν_* are the unique ergodic equilibrium states of f_* for the geometric potential $-\log |Df_*|$, so every equilibrium state is a convex combination of μ_* and ν_* [5, Corollary 3.11 and Example 3.13]. The Main Theorem applies to f_* because this map is non-renormalizable. The rate function I thus vanishes at each convex combination of μ_* and ν_* . Moreover, I can only vanish at the convex combinations of μ_* and ν_* , because the free energy function F for f_* is upper semi-continuous and therefore $I = -F$ [5,

Corollary 2.6 and Proposition 2.9]. Consider an equilibrium state μ different from the physical measure μ_* . Since $\mu \neq \mu_*$, the basin of μ has zero Lebesgue measure. Nevertheless, $I(\mu) = 0$ and therefore the LDP lower bound given by the Main Theorem show that μ does attract a significant set of initial conditions: for every $n \geq 1$ the set E_n of initial conditions x_0 for which the empirical mean $\delta_{x_0}^n$ is close to μ satisfies (2). That is, E_n is sub-exponentially large with n . Furthermore, the LDP given by the Main Theorem also shows that the equilibrium states of f_* for the potential $-\log |Df_*|$ are the only invariant measures satisfying this property. There is an analogous application of the LDP for Manneville-Pomeau maps, see [36], [7, Section 5] and [10, Appendix B]. For a certain range of parameters, there is a physical measure whose basin has full Lebesgue measure, and the rate function vanishes precisely at the convex combinations of this measure and the Dirac mass at the indifferent fixed point.

Usually the free energy function F is not upper semi-continuous,² so in general I is different from $-F$. For a concrete example for which these functions differ, consider the quadratic map $f(x) = 4x(1-x)$. Then 0 is a hyperbolic repelling fixed point and $F(\delta_0) = -\log 4$. The Lyapunov exponents of all other ergodic measures are $\log 2$, and δ_0 is weak*-approximated by measures supported on periodic points, and so $I(\delta_0) = \log 2$. For another example, consider a quadratic map f_1 given by [21, Theorem 3], whose unique physical measure is the Dirac measure supported at a repelling fixed point p of f_1 . As mentioned before $I(\delta_p) = 0$, but $F(\delta_p) = -\log |Df_1(p)| < 0$. This is also an example where the rate function vanishes at a measure that is not an equilibrium state.

In [8] a full level-2 LDP similar to the Main Theorem is shown for a positive measure set of Collet-Eckmann quadratic maps. In this result, the rate function is the same as in the Main Theorem, but instead of weighting the empirical measures with respect to the Lebesgue measure, in [8] they are measured with respect to the acip. Combining both of these LDPs, we obtain that the Lebesgue measure and the acip are sub-exponentially close on a large class of dynamically defined sets. It is not clear to us whether the LDP in [8] holds for every Collet-Eckmann quadratic map, or if a parameter exclusion as in [8] is needed.

Our methods apply with minor modifications to complex rational maps that are “backward stable” in the sense of [3, 29]; this is a condition analogous to the conclusion of Lemma 3.3. There is a large class of rational maps satisfying this property, including every polynomial with locally connected Julia set and all cycles repelling, see [29, Corollary 1]. There are however quadratic maps with all cycles repelling that are not backward stable, see [29, Remark 2]. Fur-

² Although the entropy map is upper semi-continuous as a function of measures, the Lyapunov exponent function is not lower semi-continuous in general since f has critical points, see for example [5, Proposition 2.8].

thermore, it is not known whether every rational map satisfies the specification property, or some of its consequences, like the results in [46].

1.3 Outline of the paper

In this section we outline the proof of the Main Theorem, and simultaneously describe the organization of the paper.

The proof of the Main Theorem follows the strategy originated in [7] and that has been developed in [8, 9]. The main new ingredient is a diffeomorphic pull-back argument that simplifies the construction substantially, and that allows us to apply it to a larger class of maps. The proof is divided in two parts: the lower bound is shown in Sect. 2, and the upper bound in Sects. 3 and 4.

We show that the lower bound holds without the non-flatness hypothesis. Roughly speaking, the proof of the lower bound consists of finding a set of points whose empirical means are close to a given invariant measure. In the case this last measure is hyperbolic, the desired set is easily found using Katok-Pesin theory, which allows one to approximate each hyperbolic measure by hyperbolic sets in a particular sense. The main difficulty is to deal with non-hyperbolic measures. We use the specification property to approximate a non-hyperbolic measure by hyperbolic measures, in a suitable sense. In this way we reduce the case of non-hyperbolic measures to the case of hyperbolic measures.

The upper bound is much harder, because a global control of the dynamics is required. The main idea is to construct certain horseshoes with a finite number of branches that are tailored to a given open subset of \mathcal{M} . This construction is necessarily involved due to the presence of the critical points. In [8, 9], this method was implemented under strong assumptions on the orbit of the critical value, as mentioned earlier in the introduction. In this paper, we use a diffeomorphic pull-back argument to replace the analytic horseshoe constructions in [8, 9] by one of more topological flavor, enabling us to dispense with the strong assumptions on the critical orbits altogether.

The diffeomorphic pull-back argument is developed in Sect. 3, where it is stated as the “Uniform Scale Lemma.” One of the main ingredients in the proof of this lemma are some general sub-exponential distortion bounds (Proposition 3.1 in Sect. 3.1.) These sub-exponential distortion bounds are combined with a method that goes back to [39], to carefully avoid critical points and choose diffeomorphic pull-backs. The preliminary results needed to implement this method are established in Sect. 3.2, and the proof of the Uniform Scale Lemma is given in Sect. 3.3.

The proof of the upper bound is completed in Sect. 4. The main step is to construct, for a given basic open set of $\mathcal{M}(f)$ and for each large integer $n \geq 1$, a certain horseshoe with inducing time q , where $q \geq n$ and $q - n = o(n)$ as $n \rightarrow \infty$ (Proposition 4.1 in Sect. 4.1.) By a *horseshoe with inducing*

time q we mean a finite collection L_1, L_2, \dots, L_t of pairwise disjoint closed intervals such that f^q maps each $L_i, i \in \{1, 2, \dots, t\}$, diffeomorphically onto an interval whose interior contains $\bigcup_{i=1}^t L_i$. The inducing time q consists of three explicit parts: in the first n iterations, the intervals are mapped to a ball of radius $n^{-\alpha}$, for a fixed constant $\alpha > 1$, centered at a carefully chosen base point; in the second part, of roughly $\log n$ iterations, intervals reach a fixed scale $\kappa > 0$ independent of n ; the third part, of a bounded number of iterations, the intervals return to a prefixed small interval. In order to reach the scale κ , a key ingredient is the Uniform Scale Lemma in Sect. 3. Once the horseshoe is constructed, we prove two intermediate estimates in Sect. 4.2. The first is restricted to a small interval (Proposition 4.4), and the second is a global estimate (Proposition 4.6) obtained by using topological exactness to spread out the local estimate. The local estimate is used to treat inflection critical points. The proof of the upper bound is completed in Sect. 4.3.

1.4 Notation

The following notation and terms are used in the rest of the paper. For $x \in X$ and $\eta > 0$ denote by $B(x, \eta)$ the closed ball of radius η centered at x , *i.e.*,

$$B(x, \eta) = \{y \in X : |y - x| \leq \eta\},$$

and for subsets A and A' of X define

$$B(A, \eta) = \bigcup_{x \in A} B(x, \eta), \quad \text{dist}(x, A) = \inf\{|x - a| : a \in A\},$$

and

$$\text{dist}(A, A') = \inf\{|a - a'| : a \in A, a' \in A'\}.$$

A subset F of X is called η -dense if $B(F, \eta) = X$ holds. For a subset A of X , denote by $\text{HD}(A)$ the Hausdorff dimension of A .

Let $f : X \rightarrow X$ be continuously differentiable. A subset K of X is *forward f -invariant* if $f(K) \subset K$. The set K is called *hyperbolic*, if there exist $C > 0$ and $\lambda > 1$ such that for every $x \in K$ and every integer $n \geq 1$, $|Df^n(x)| \geq C\lambda^n$ holds.

2 Large deviations lower bound

In this section we prove the large deviations lower bound in the Main Theorem. As the proof below shows, these estimates hold without the non-flatness

hypothesis. The following is the key estimate and it contains Ruelle's inequality. It must be noted that in the following estimate we have to treat measures with zero Lyapunov exponent.

Proposition 2.1 (Key Estimate) *Let $f: X \rightarrow X$ have Hölder continuous derivative and at most a finite number of critical points. Assume f is topologically exact. Let $l \geq 1$ be an integer, $\varphi_1, \varphi_2, \dots, \varphi_l: X \rightarrow \mathbb{R}$ continuous functions and $\alpha_1, \alpha_2, \dots, \alpha_l \in \mathbb{R}$. Then for every $\mu \in \mathcal{M}(f)$ such that $\int \varphi_j d\mu > \alpha_j$ for every $j \in \{1, 2, \dots, l\}$,*

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \left| \left\{ x \in X : \frac{1}{n} S_n \varphi_j(x) > \alpha_j \text{ for every } j \in \{1, 2, \dots, l\} \right\} \right| \geq F(\mu).$$

In the proof of this proposition we use the following version of Katok's theorem, which allows one to approximate each hyperbolic measure by hyperbolic sets in a particular sense, compare with [24, Theorem S.5.9] and [38, Theorem 4.1]. Using Dobbs' adaptation of Pesin's theory to interval maps [16, Theorem 6], the proof is a slight modification of that of [24, Theorem S.5.9] and hence we omit it. For a continuous map $f: X \rightarrow X$, a subset U of X , and an integer $n \geq 1$, each connected component of $f^{-n}(U)$ is called a *pull-back* of U by f^n . If in addition f is differentiable, then a pull-back J of U by f^n is called *diffeomorphic* if $f^n: J \rightarrow U$ is a diffeomorphism.

Lemma 2.2 *Let $f: X \rightarrow X$ have Hölder continuous derivative and at most a finite number of critical points. Let $\mu \in \mathcal{M}(f)$ be ergodic and such that $h(\mu) > 0$. Let $l \geq 1$ be an integer, and $\varphi_1, \dots, \varphi_l: X \rightarrow \mathbb{R}$ continuous functions. Then for every $\varepsilon > 0$ there are integers $k \geq 2$ and $m \geq 1$ satisfying $\frac{1}{m} \log k \geq h(\mu) - \varepsilon$, a closed subinterval K of X , and pairwise disjoint diffeomorphic pull-backs K_1, \dots, K_k of K by f^m contained in K , such that the following holds:*

$$\left| \frac{1}{m} S_m \varphi_j(x) - \int \varphi_j d\mu \right| < \varepsilon \text{ for every } x \in \bigcup_{i=1}^k K_i \text{ and every } j \in \{1, \dots, l\};$$

and

$$e^{(\lambda(\mu)-\varepsilon)m} \leq |Df^m(x)| \leq e^{(\lambda(\mu)+\varepsilon)m} \text{ for every } x \in \bigcup_{i=1}^k K_i.$$

Proof of Proposition 2.1 Fix $\varepsilon > 0$ sufficiently small so that $\int \varphi_j d\mu > \alpha_j + \varepsilon$ holds for every $j \in \{1, \dots, l\}$. For each $(n_0, \dots, n_{l+1}) \in \mathbb{Z}^{l+2}$ put

$$C((n_0, \dots, n_{l+1})) = \left[n_0 \frac{\varepsilon}{3}, (n_0 + 1) \frac{\varepsilon}{3} \right) \times \dots \times \left[n_{l+1} \frac{\varepsilon}{3}, (n_{l+1} + 1) \frac{\varepsilon}{3} \right).$$

Denote by $\mathcal{M}_{\text{erg}}(f)$ the subset of $\mathcal{M}(f)$ of ergodic measures, and let $\Phi: \mathcal{M}_{\text{erg}}(f) \rightarrow \mathbb{R}^{l+2}$ be the function defined by

$$\Phi(\nu) = \left(h(\nu), \chi(\nu), \int \varphi_1 d\nu, \dots, \int \varphi_l d\nu \right).$$

Finally, let Z be the subset of \mathbb{Z}^{l+2} of those \underline{n} such that $\Phi^{-1}(C(\underline{n}))$ is nonempty, set $s = \#Z$, choose a bijection $\iota: \{1, \dots, s\} \rightarrow Z$, and for each $i \in \{1, \dots, s\}$ choose a measure $\mu_i \in \mathcal{M}_{\text{erg}}(f)$ in $\Phi^{-1}(C(\iota(i)))$. Thus, if $\underline{\mu}$ is the unique probability measure on $\mathcal{M}_{\text{erg}}(f)$ such that $\mu = \int \nu d\underline{\mu}(\nu)$, and for each $i \in \{1, \dots, s\}$ we put $\beta_i = \underline{\mu}(\Phi^{-1}(C(\iota(i))))$, then the measure $\mu' = \beta_1 \mu_1 + \dots + \beta_s \mu_s$ is in $\mathcal{M}(f)$, and satisfies $|h(\mu) - h(\mu')| \leq \frac{\varepsilon}{3}$, $|\lambda(\mu) - \lambda(\mu')| \leq \frac{\varepsilon}{3}$, and for each $j \in \{1, \dots, l\}$,

$$\left| \int \varphi_j d\mu - \int \varphi_j d\mu' \right| \leq \frac{\varepsilon}{3}.$$

For each $i \in \{1, \dots, s\}$ define integers k_i and m_i and subintervals $K^i, K_1^i, \dots, K_{k_i}^i$ of X , as follows. In the case where $h(\mu_i) > 0$, let $k_i = k, m_i = m, K^i = K, K_1^i = K_1, \dots, K_{k_i}^i = K_{k_i}$ be as in Lemma 2.2 with ε replaced by $\frac{\varepsilon}{3}$. Suppose $h(\mu_i) = 0$. By [46, Theorem 1] and the upper semi-continuity of the Lyapunov exponent function there is a periodic point p such that, if we denote by $N \geq 1$ its minimal period, then $\frac{1}{N} \log |Df^N(p)| \leq \lambda(\mu_i) + \frac{\varepsilon}{6}$ and for each $j \in \{1, \dots, l\}$,

$$\left| \frac{1}{N} S_N \varphi_j(p) - \int \varphi_j d\mu_i \right| \leq \frac{\varepsilon}{6}.$$

Using that f is topologically exact, it follows that for every sufficiently small interval K containing p , the pull-back K_1 of K by f^N containing p is contained in K . Reduce K if necessary, so that $f^N: K_1 \rightarrow K$ is a diffeomorphism, and such that for every $x \in K_1$ we have $\frac{1}{N} \log |Df^N(x)| \leq \lambda(\mu_i) + \frac{\varepsilon}{3}$ and for each $j \in \{1, \dots, l\}$,

$$\left| \frac{1}{N} S_N \varphi_j(x) - \int \varphi_j d\mu_i \right| \leq \frac{\varepsilon}{3}.$$

Set $k_i = 1, m_i = N, K^i = K$, and $K_1^i = K_1$.

Take an integer $M \geq 1$ such that for each $i \in \{1, \dots, s\}$ we have $f^M(K^i) = X$, and fix an integer $n \geq 1$. For each $i \in \{1, \dots, s\}$, put

$$\ell_i = \left\lceil \frac{\beta_i n}{m_i} \right\rceil \text{ and } n_i = \ell_i m_i + M,$$

and denote by \mathcal{L}_i the collection connected components of $\left(f^{m_i}|_{K_1^i \cup \dots \cup K_{s(i)}^i}\right)^{-\ell_i}(K^i)$. Note that $\#\mathcal{L}_i = k_i^{\ell_i}$, and that for each $L \in \mathcal{L}_i$ we have $f^{n_i}(L) = X$. Furthermore, for each $x \in L$ we have

$$\frac{1}{n_i} \log |Df^{n_i}(x)| \leq \frac{\ell_i m_i}{n_i} \left(\lambda(\mu_i) + \frac{\varepsilon}{3} \right) + \frac{M}{n_i} \log \left(\sup_X |Df| \right), \quad (3)$$

and for each $j \in \{1, \dots, l\}$ we have

$$\left| \frac{1}{n_i} S_{n_i} \varphi_j(x) - \int \varphi_j d\mu_i \right| \leq \frac{\ell_i m_i}{n_i} \frac{\varepsilon}{3} + \frac{M}{n_i} \sup_X |\varphi_j|. \quad (4)$$

Set $m = n_1 + \dots + n_s$, and note that the sets in

$$\mathcal{L} = \left\{ (f^{n_1}|_{L_1})^{-1} \circ \dots \circ (f^{n_s}|_{L_s})^{-1}(X) : L_1 \in \mathcal{L}_1, \dots, L_s \in \mathcal{L}_s \right\}$$

are pairwise disjoint, and that each set in \mathcal{L} is mapped onto X by f^m . On the other hand, if n is sufficiently large, then

$$\begin{aligned} \frac{1}{m} \log(\#\mathcal{L}) &= \frac{\ell_1 \log k_1 + \dots + \ell_s \log k_s}{n_1 + \dots + n_s} \geq \left(\sum_{i=1}^s \frac{\beta_i}{m_i} \log k_i \right) - \frac{\varepsilon}{3} \\ &\geq \left(\sum_{i=1}^s \beta_i \left(h(\mu_i) - \frac{\varepsilon}{3} \right) \right) - \frac{\varepsilon}{3} = h(\mu') - \frac{2}{3}\varepsilon \geq h(\mu) - \varepsilon. \end{aligned}$$

Furthermore, by (3), for each $L \in \mathcal{L}$ and $x \in L$ we have $|Df^m(x)| \leq e^{(\lambda(\mu)+\varepsilon)m}$, and by (4), for each $j \in \{1, \dots, l\}$ we have

$$\left| \frac{1}{m} S_m \varphi_j(x) - \int \varphi_j d\mu \right| \leq \varepsilon.$$

Note that for each L in \mathcal{L} we have $|L| \geq e^{-(\lambda(\mu)+\varepsilon)m}$.

Let n be a large integer and write $n = pm + q$, where p, q are non-negative integers with $0 \leq q \leq m - 1$. We have

$$\begin{aligned}
& \frac{1}{n} \log \left| \left\{ x \in X : \frac{1}{n} S_n \varphi_j(x) > \alpha_j \text{ for every } j \in \{1, \dots, l\} \right\} \right| \\
& \geq \frac{1}{n} \log \left| \left\{ x \in X : \frac{1}{pm} S_{pm} \varphi_j(x) > \alpha_j + \varepsilon \text{ for every } j \in \{1, \dots, l\} \right\} \right| \\
& \geq \frac{1}{n} \log \left(\sum_{L \in \bigvee_{i=0}^{p-1} f^{-im} \mathcal{L}} |L| \right) \\
& \geq \frac{1}{n} (p \log(\# \mathcal{L}) - pm(\lambda(\mu) + \varepsilon)) \\
& \geq \frac{1}{m} \log(\# \mathcal{L}) - (\lambda(\mu) + 2\varepsilon) \\
& \geq h(\mu) - \lambda(\mu) - 3\varepsilon.
\end{aligned}$$

Letting $n \rightarrow \infty$ and then $\varepsilon \rightarrow 0$ we obtain the desired inequality. \square

Proof of the large deviations lower bound in the Main Theorem Let $f: X \rightarrow X$ be a map satisfying the hypotheses of Proposition 2.1, and \mathcal{G} an open subset of \mathcal{M} . Note that the topology of \mathcal{M} has a base consisting of sets of the form

$$\left\{ \nu \in \mathcal{M} : \int \varphi_j d\nu > \alpha_j \text{ for every } j \in \{1, \dots, l\} \right\},$$

where $l \geq 1$ is an integer, each $\varphi_j: X \rightarrow \mathbb{R}$ is a continuous function and $\alpha_j \in \mathbb{R}$. Hence, there exists a collection $\{\mathcal{O}_\xi\}_\xi$ of sets of this form such that $\mathcal{G} = \bigcup_\xi \mathcal{O}_\xi$. Proposition 2.1 applied to each \mathcal{O}_ξ yields

$$\begin{aligned}
\liminf_{n \rightarrow \infty} \frac{1}{n} \log |\{x \in X : \delta_x^n \in \mathcal{G}\}| &= \liminf_{n \rightarrow \infty} \frac{1}{n} \log \left| \left\{ x \in X : \delta_x^n \in \bigcup_\xi \mathcal{O}_\xi \right\} \right| \\
&\geq \sup_\xi \liminf_{n \rightarrow \infty} \frac{1}{n} \log |\{x \in X : \delta_x^n \in \mathcal{O}_\xi\}| \\
&\geq \sup_\xi \sup_{\mathcal{O}_\xi} F \\
&= \sup_{\mathcal{G}} F \\
&= -\inf_{\mathcal{G}} I.
\end{aligned}$$

\square

3 The uniform scale lemma

This section is devoted to the proof of the following lemma that is a key element of the proof of the large deviations upper bound in the Main Theorem. The large deviations upper bound is completed in Sect. 4.

For a differentiable map $g: X \rightarrow X$ and a subinterval J of X that does not contain critical points of g , the *distortion of g on J* is by definition

$$\sup \left\{ \frac{|Dg(x)|}{|Dg(y)|} : x, y \in J \right\}.$$

Uniform Scale Lemma *Let $f: X \rightarrow X$ have Hölder continuous derivative and only non-flat critical points. Assume f is topologically exact. Then for every $\varepsilon > 0$ there exist constants $\eta_0 > 0$, $C > 0$, and $\kappa > 0$, such that for every $\eta \in (0, \eta_0)$ there is $n_0 \geq 1$ such that the following property holds for every integer $n \geq n_0$. For every subinterval W of X that satisfies $\eta \leq |f^n(W)| \leq 2\eta$, there exists a subinterval J of W and an integer m such that*

$$|J| \geq e^{-\varepsilon n} |W|, \quad n \leq m \leq n + C \log n, \quad |f^m(J)| \geq \kappa,$$

and such that f^m maps J diffeomorphically onto $f^m(J)$ with distortion bounded by $e^{\varepsilon n}$ (Fig. 1).

In Sect. 3.1 we establish one of the main ingredients in the proof of this lemma, which are some general sub-exponential distortion bounds (Proposition 3.1). The first type of distortion bound is on the ratio of the sizes of two iterated intervals, which holds for an arbitrary pull-back that is not necessarily diffeomorphic. The second one is a sub-exponential derivative distortion bound for diffeomorphic pull-backs with a definite ‘‘Koebe space’’. This last distortion bound is obtained from the Koebe Principle in [13] and a sub-exponential cross-ratio distortion bound. In Sect. 3.2 we show the abundance of ‘‘safe points’’ contained in hyperbolic sets (Lemma 3.5). This is used to apply the method of [39] to find sub-exponentially small intervals all whose pull-backs by a high iterate of the map are mapped diffeomorphically to unit scale. The proof of the Uniform Scale Lemma is given in Sect. 3.3.

3.1 Sub-exponential distortion bounds

In this section we prove the following proposition giving a sub-exponential bound on the ratio of the sizes of two iterated intervals, and a sub-exponential derivative distortion bound for certain diffeomorphic pull-backs.

Proposition 3.1 *Let $f: X \rightarrow X$ have Hölder continuous derivative and only non-flat critical points. Assume f is topologically exact. Then for every $\varepsilon > 0$*

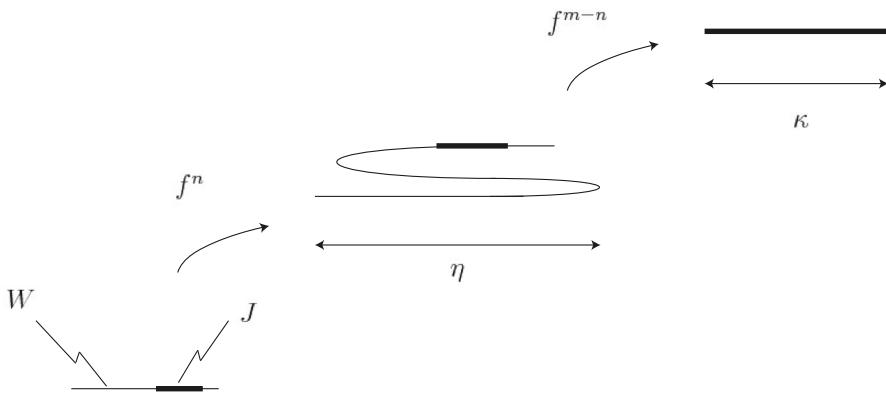


Fig. 1 On the Uniform Scale Lemma: for a given $\varepsilon > 0$ one can find two small scales $\eta > 0$ and $\kappa > 0$ such that for every pull-back W of intervals of size η one can choose a subinterval J of W that is mapped diffeomorphically to an interval of length κ in time $m, n \leq m \leq n + C \log n$

there exist an integer $n_1 \geq 1$ and $\eta_0 > 0$ such that for every integer $n \geq n_1$, every subinterval W of X that satisfies $|f^n(W)| \leq 2\eta_0$, and for every subinterval J of W ,

$$\frac{|f^n(J)|}{|f^n(W)|} \leq e^{\varepsilon n} \frac{|J|}{|W|}.$$

If in addition $f^n: W \rightarrow f^n(W)$ is a diffeomorphism and $|f^n(J)| \leq \text{dist}(\partial f^n(W), f^n(J))$, then the distortion of f^n on J is bounded by $e^{\varepsilon n}$.

For the proof of this proposition we need the next lemma, in which we use the assumption that each critical point is non-flat. To state this lemma, we use the concept of “cross-ratio” that we proceed to recall. Given a subinterval \widehat{J} of \mathbb{R} and an interval J whose closure is contained in the interior of \widehat{J} , denote by L and R the connected components of $\widehat{J} \setminus J$. Then the *cross-ratio* $\text{Cr}(\widehat{J}; J)$ of \widehat{J} and J is defined by

$$\text{Cr}(\widehat{J}; J) = \frac{|\widehat{J}| |J|}{|L| |R|}.$$

Lemma 3.2 *Let $f: X \rightarrow X$ be continuously differentiable with only non-flat critical points. Then there exist constants $C_0 > 1$ and $\delta_0 > 0$ such that for every interval \widehat{U} contained in $B(\text{Crit}(f), \delta_0)$, and every subinterval U of \widehat{U} ,*

$$\frac{|f(U)|}{|f(\widehat{U})|} \leq C_0 \frac{|U|}{|\widehat{U}|}.$$

If in addition \widehat{U} is disjoint from $\text{Crit}(f)$ and the closure of U is contained in the interior of \widehat{U} , then

$$\text{Cr}(f(\widehat{U}); f(U)) \geq C_0^{-1} \text{Cr}(\widehat{U}; U).$$

Proof Let $c \in \text{Crit}(f)$. By the definition of non-flatness, there exist a number $\ell > 1$ and diffeomorphisms ϕ and ψ of \mathbb{R} such that $\phi(c) = \psi(f(c)) = 0$ and $g = \psi \circ f \circ \phi^{-1}$ satisfies $|g(x)| = |x|^\ell$ for x near 0. It is thus enough to prove the lemma with f replaced by g . For g , the second inequality with $C_0 = 1$ is given by [13, Property 4 in Sect.IV.1] by noting that the Schwarzian derivative of g is negative on $\mathbb{R} \setminus \{0\}$. To prove the first inequality we treat four cases separately.

Case 1: $0 \in U$. We have $(|U|/2)^\ell \leq |g(U)| \leq |U|^\ell$. Since $0 \in \widehat{U}$ we also have $(|\widehat{U}|/2)^\ell \leq |g(\widehat{U})| \leq |\widehat{U}|^\ell$. Then $|g(U)|/|g(\widehat{U})| \leq (2|U|/|\widehat{U}|)^\ell < 2^\ell |U|/|\widehat{U}|$.

Case 2: $0 \notin U$ and $0 \in \widehat{U}$. By the mean value theorem and the form of g , there is ξ in U such that $|g(U)| = |Dg(\xi)| \cdot |U| \leq \ell |\widehat{U}|^{\ell-1} \cdot |U|$. Combining this with the lower estimate of $|g(\widehat{U})|$ in Case 1 yields $|g(U)|/|g(\widehat{U})| \leq 2^\ell \ell |U|/|\widehat{U}|$.

Case 3: $0 \notin \widehat{U}$ and $|\widehat{U}| \leq \text{dist}(0, \widehat{U})$. The mean value theorem gives $|g(U)| = |Dg(\xi)| \cdot |U|$ and $|g(\widehat{U})| = |Dg(\eta)| \cdot |\widehat{U}|$ for some $\xi \in U$ and $\eta \in \widehat{U}$. The assumption $|\widehat{U}| \leq \text{dist}(0, \widehat{U})$ implies $|\xi/\eta| \leq 2$, and so $|g(U)|/|g(\widehat{U})| \leq 2^{\ell-1} |U|/|\widehat{U}|$.

Case 4: $0 \notin \widehat{U}$ and $|\widehat{U}| > \text{dist}(0, \widehat{U})$. Let V denote the smallest closed interval containing \widehat{U} and 0. We have $|g(V)| = |g(\widehat{U})| + |g(V \setminus \widehat{U})| < 2|g(\widehat{U})|$. Using this and the estimate in Case 2 for the pair (U, V) yields

$$|g(U)|/|g(\widehat{U})| < (1/2)|g(U)|/|g(V)| \leq 2^{\ell-1} \ell |U|/|V| < 2^{\ell-1} \ell |U|/|\widehat{U}|.$$

□

In the proof of Proposition 3.1 we also use general properties of topologically exact maps. First, notice that from the compactness of X , for every continuous and topologically exact map $f: X \rightarrow X$ and each $\gamma > 0$ there is an integer $N \geq 1$ such that for every subinterval J of X with $|J| \geq \gamma$, we have $f^N(J) = X$; we denote by $N(\gamma)$ the smallest such integer.

Lemma 3.3 *Let $f: X \rightarrow X$ be a continuous map that is topologically exact. Then for every $\varepsilon > 0$ there exists $\eta \in (0, 1/2)$ such that for every integer $n \geq 1$ and every subinterval W of X that satisfies $|f^n(W)| \leq \eta$, $|f^i(W)| \leq \varepsilon$ holds for every $i \in \{0, \dots, n-1\}$.*

Proof Let $\eta \in (0, 1/2)$ be such that for every subinterval V of X that satisfies $|V| \leq \eta$, $|f^i(V)| \leq 1/2$ holds for every $i \in \{0, \dots, N(\varepsilon) - 1\}$. Let $n \geq 1$ be an integer and W a subinterval of X such that $|f^n(W)| \leq \eta$. If

$|f^{i_0}(W)| > \varepsilon$ holds for some $i_0 \in \{0, \dots, n-1\}$, then the definition of $N(\varepsilon)$ gives $f^{N(\varepsilon)}(f^{i_0}(W)) = X$. Since $f(X) = X$ we get $f^{N(\varepsilon)-1}(f^n(W)) = X$, and this contradicts the choice of η with $V = f^n(W)$. \square

Proof of Proposition 3.1 In order to treat critical relations that can arise in the case $\# \text{Crit}(f) \geq 2$ we introduce the following notion. We say $c \in \text{Crit}(f)$ is a *tail* if $f^n(c) \notin \text{Crit}(f)$ holds for every $n \geq 1$. Let $\text{Crit}'(f)$ denote the set of tails.

Consider a graph made up of vertices and oriented edges between them. The vertices are critical points of f . For two vertices c_0 and c_1 put an edge from c_0 to c_1 if there exists an integer $n \geq 1$ such that $f(c_0), f^2(c_0), \dots, f^{n-1}(c_0) \notin \text{Crit}(f)$ and $f^n(c_0) = c_1$. The edge is labeled with n . By definition, there is at most one outgoing edge from each vertex. Since no critical point is periodic, there is no loop in the graph. The concatenation of edges groups the set of vertices into blocks, which might intersect. For each block consider the sum of labels of all its edges. Let E denote the maximal sum over all blocks. Let $\varepsilon > 0$ be given and let C_0 and δ_0 be the constants given by Lemma 3.2. Choose a sufficiently large integer $n_1 \geq 1$ such that $e^{\varepsilon n_1/12} \geq 2C_0^{2E}$. Let $\delta \in (0, \delta_0)$ be such that the set $\bigcup_{j=1}^{n_1} f^j(B(\text{Crit}'(f), \delta))$ is disjoint from $B(\text{Crit}(f), \delta/2)$.

Since f is continuously differentiable, there is $\kappa \in (0, \delta/2)$ such that for every interval U contained in $X \setminus B(\text{Crit}(f), \delta/2)$ that satisfies $|U| \leq \kappa$,

$$\sup_{x, y \in U} \frac{|Df(x)|}{|Df(y)|} \leq e^{\frac{\varepsilon}{24}}. \quad (5)$$

Finally, in view of Lemma 3.3 we can choose $\eta_0 > 0$ such that for every $\eta \in (0, \eta_0)$, every $x \in X$, every integer $n \geq 1$ and every pull-back W of $B(x, \eta)$ by f^n , $|f^j(W)| \leq \kappa$ holds for every $j \in \{0, \dots, n-1\}$. Note that by our choices of n_1 and δ , it follows that

$$\#\{j \in \{0, \dots, n-1\} : f^j(W) \cap B(\text{Crit}(f), \delta/2) \neq \emptyset\} \leq E \left(\frac{n}{n_1} + 1 \right) \leq \frac{2En}{n_1}. \quad (6)$$

Let $n \geq n_1$, $\eta \in (0, \eta_0)$, W a pull-back of $B(x, \eta)$ by f^n and J a subinterval of W . For every $j \in \{0, \dots, n-1\}$ we have $|f^j(W)| \leq \kappa$. Thus, if in addition $f^j(W)$ is disjoint from $B(\text{Crit}(f), \delta/2)$, then (5) gives

$$\frac{|f^{j+1}(J)|}{|f^{j+1}(W)|} \leq e^{\frac{\varepsilon}{24}} \frac{|f^j(J)|}{|f^j(W)|}.$$

If in addition $f^j(W)$ is disjoint from $\text{Crit}(f)$, then for every subinterval \widehat{U} of $f^j(W)$ and every interval U whose closure is contained in the interior of \widehat{U} ,

$$\text{Cr}(f^{j+1}(\widehat{U}); f^{j+1}(U)) \geq e^{-\frac{\varepsilon}{12}} \text{Cr}(f^j(\widehat{U}); f^j(U)).$$

Suppose now $j \in \{0, \dots, n-1\}$ is such that $f^j(W)$ intersects $B(\text{Crit}(f), \delta/2)$. Since $\kappa \in (0, \delta/2)$, the interval $f^j(W)$ is contained in $B(\text{Crit}(f), \delta)$, and by Lemma 3.2 we have

$$\frac{|f^{j+1}(J)|}{|f^{j+1}(W)|} \leq C_0 \frac{|f^j(J)|}{|f^j(W)|}.$$

If in addition $f^j(W)$ is disjoint from $\text{Crit}(f)$, then for every subinterval \widehat{U} of $f^j(W)$ and every interval U whose closure is contained in the interior of \widehat{U} ,

$$\text{Cr}(f^{j+1}(\widehat{U}); f^{j+1}(U)) \geq C_0^{-1} \text{Cr}(f^j(\widehat{U}); f^j(U)).$$

Therefore, by our choice of n_1 and (6) we have

$$\frac{|f^n(J)|}{|f^n(W)|} \leq C_0^{\frac{2En}{n_1}} e^{\frac{\varepsilon}{12}n} \leq e^{\varepsilon n} \frac{|J|}{|W|},$$

which gives the first assertion of the proposition.

To prove the second assertion of the proposition, suppose $f^n: W \rightarrow f^n(W)$ is a diffeomorphism. Then for every subinterval \widehat{U} of W and interval U whose closure is contained in the interior of \widehat{U} ,

$$\frac{\text{Cr}(f^n(\widehat{U}); f^n(U))}{\text{Cr}(\widehat{U}; U)} = \prod_{j=0}^{n-1} \frac{\text{Cr}(f^{j+1}(\widehat{U}); f^{j+1}(U))}{\text{Cr}(f^j(\widehat{U}); f^j(U))} \geq C_0^{-\frac{2En}{n_1}} e^{-\frac{\varepsilon}{12}n} \geq 2e^{-\frac{\varepsilon}{6}n}.$$

The Koebe Principle [13, Theorem IV.1.2] with $\tau = 1$ implies that the distortion of f^n on J is bounded by $e^{\varepsilon n}$. This completes the proof of the proposition. \square

3.2 Abundance of safe points in hyperbolic sets

Let $f: X \rightarrow X$ be a differentiable interval map with at most a finite number of critical points. In order to carefully avoid critical points and choose diffeomorphic pull-backs, we use the method introduced in [39]. We adopt the terminology of “safe points” in [40, Definition 12.5.7]. For a given $\alpha > 0$ and an integer $n \geq 1$ define

$$E_n(\alpha) = \bigcup_{j=1}^{\infty} B(f^j(\text{Crit}(f)), \min\{n^{-\alpha}, j^{-\alpha}\}).$$

Note that the set $E_n(\alpha)$ is decreasing in n . Set

$$E(\alpha) = \bigcap_{n=1}^{\infty} E_n(\alpha).$$

Note that $E(\alpha)$ contains $\bigcup_{j=1}^{\infty} f^j(\text{Crit}(f))$.

We say $x \in X$ is α -safe if $x \notin E(\alpha)$. If x is α -safe, then for every integer $n \geq 1$ with $x \notin E_n(\alpha)$ the ball $B(x, n^{-\alpha})$ is disjoint from $\bigcup_{j=1}^n f^j(\text{Crit}(f))$. Hence, the pull-backs of $B(x, n^{-\alpha})$ by f^n are diffeomorphic.

Lemma 3.4 *For every $\alpha > 0$, $\text{HD}(E(\alpha)) \leq \alpha^{-1}$.*

Proof For each n consider the covering of $E(\alpha)$ by the intervals

$$B(f^j(c), \min\{n^{-\alpha}, j^{-\alpha}\}), \quad c \in \text{Crit}(f), \quad j \in \{1, 2, \dots\}.$$

Let $\beta > \alpha^{-1}$. We have

$$\begin{aligned} & \sum_{c \in \text{Crit}(f)} \sum_{j=1}^{\infty} |B(f^j(c), \min\{n^{-\alpha}, j^{-\alpha}\})|^{\beta} \\ &= \sum_{c \in \text{Crit}(f)} \left(\sum_{j=1}^n + \sum_{j=n+1}^{\infty} \right) \\ &\leq \# \text{Crit}(f) \cdot \left(2^{\beta} n^{1-\alpha\beta} + \sum_{j=n+1}^{\infty} 2^{\beta} j^{-\alpha\beta} \right). \end{aligned}$$

This number goes to 0 as $n \rightarrow \infty$, and so the Hausdorff β -measure of $E(\alpha)$ is 0. Since $\beta > \alpha^{-1}$ is arbitrary we obtain $\text{HD}(E(\alpha)) \leq \alpha^{-1}$. \square

Lemma 3.5 *Let $f : X \rightarrow X$ have Hölder continuous derivative and at most a finite number of critical points. If f is topologically exact, then there is $\alpha > 0$ such that the following property holds. For every $\eta > 0$ there is a hyperbolic set Λ of f such that for every $x \in X$, the set $B(x, \eta) \cap \Lambda$ is nonempty and contains an α -safe point.*

Proof Since f is topologically exact, there exist an integer $n > 0$ and a closed subset \widehat{A} of X such that $f^n(\widehat{A}) \subset \widehat{A}$ and $f^n : \widehat{A} \rightarrow f^n(\widehat{A})$ is topologically conjugate to the one-sided full shift on two symbols. Hence, f has positive topological entropy, see also [45, Proposition 4.70]. From the variational principle, see for example [25, Theorem 4.4.11] or [40, Theorem 3.4.1], there is a measure μ in $\mathcal{M}(f)$ satisfying $h(\mu) > 0$, and therefore $\lambda(\mu) > 0$ by

Ruelle's inequality. By Lemma 2.2 with $\varepsilon = \lambda(\mu)/2$, there are integers $k \geq 2$ and $m \geq 1$, a closed subinterval K of X and pairwise disjoint closed subintervals K_1, \dots, K_k of K , such that for each i in $\{1, \dots, k\}$ the map $f^m: K_i \rightarrow K$ is a diffeomorphism and $|Df^m| \geq \exp(\lambda(\mu)m/2)$ on K_i . It follows that the maximal invariant set $\widehat{\Lambda}_0$ of f^m on $\bigcup_{i=1}^k K_i$ is a hyperbolic set for f^m . Since $k \geq 2$, we have $\text{HD}(\widehat{\Lambda}_0) > 0$.

Let $Q \geq 2\eta^{-1}$ be an integer and put $\xi = \exp(\lambda(\mu)m/2)$. Since f is topologically exact, the map f^m is also topologically exact, so there is an integer $N \geq 1$ such that $f^{Nm}\left(\left(\frac{i-1}{Q}, \frac{i}{Q}\right)\right) = X$ holds for each $i \in \{1, \dots, Q\}$. Let p_0 be a point in the uncountable set $\widehat{\Lambda}_0$ that is not in $\bigcup_{j=1}^{\infty} f^j(\text{Crit}(f))$. Define recursively for each $i \in \{1, \dots, Q\}$ a point $p_i \in \left(\frac{i-1}{Q}, \frac{i}{Q}\right)$, so that $f^{Nm}(p_i) = p_{i-1}$. Using again that f^m is topologically exact, we can find an integer $N' \geq 1$ and a point p in the interior of K that is not in $\widehat{\Lambda}_0$, such that $f^{N'm}(p) = p_Q$. Defining $\ell = QN + N'$, we have that $f^{\ell m}(p) = p_0$ and that the set

$$\{p, f^m(p), \dots, f^{\ell m}(p)\} \supset \{p_1, p_2, \dots, p_Q\}$$

is η -dense in X . Since p_0 is not in $\bigcup_{j=1}^{\infty} f^j(\text{Crit}(f))$, there is $\delta_0 > 0$ such that $B(p_0, \delta_0)$ is disjoint from $\bigcup_{j=1}^{\ell m} f^j(\text{Crit}(f))$. It follows that the pull-back W_0 of $B(p_0, \delta_0)$ by $f^{\ell m}$ containing p is diffeomorphic. Reduce δ_0 if necessary so that W_0 is contained in K . Let $\ell_0 \geq 1$ be a sufficiently large integer such that $\xi^{-\ell_0} < \inf_{W_0} |Df^{\ell_0 m}|$ and such that the pull-back of K by $f^{\ell_0 m}$ containing p_0 is contained in $B(p_0, \delta_0)$. Since p_0 is in $\widehat{\Lambda}_0$, it follows that this last pull-back is diffeomorphic. We conclude that, if we put $M = (\ell + \ell_0)m$, then the pull-back L_0 of K by f^M containing p is diffeomorphic. Moreover, from our choice of ℓ_0 we have

$$\inf_{L_0} |Df^M| \geq \xi^{\ell_0} \inf_{W_0} |Df^{\ell_0 m}| > 1. \quad (7)$$

Let \mathcal{L} be the collection formed by L_0 and by all pull-backs of K by f^M that intersect $\widehat{\Lambda}_0$. Since $\inf_{K_i} |Df^m| \geq \xi$ for each $i \in \{1, \dots, k\}$, $\inf_{L \in \mathcal{L}} |Df^M| \geq \xi^{\ell + \ell_0} > 1$ holds for every $L \in \mathcal{L}$ different from L_0 . Together with (7) this implies that the maximal invariant set $\widehat{\Lambda}$ of f^M in $\bigcup_{L \in \mathcal{L}} L$ is a hyperbolic set for f^M , and that $f^M: \widehat{\Lambda} \rightarrow \widehat{\Lambda}$ is topologically exact. On the other hand, the point p is by definition in L_0 and $f^M(p) = f^{\ell_0 m}(p_0)$ is in $\widehat{\Lambda}_0$. This implies $p \in \widehat{\Lambda}$ and therefore $\widehat{\Lambda}$ is η -dense on X . So, for every $x \in X$ the ball $B(x, \eta)$ intersects $\widehat{\Lambda}$ and, since $f^M: \widehat{\Lambda} \rightarrow \widehat{\Lambda}$ is topologically exact, it follows that there is an integer $k \geq 1$ such that $f^{kM}(B(x, \eta) \cap \widehat{\Lambda}) = \widehat{\Lambda}$. Using that f^{kM} is

Lipschitz continuous on $\widehat{\Lambda}$ and that $\widehat{\Lambda}$ contains $\widehat{\Lambda}_0$, we obtain

$$\text{HD}(B(x, \eta) \cap \widehat{\Lambda}) \geq \text{HD}(\widehat{\Lambda}) \geq \text{HD}(\widehat{\Lambda}_0).$$

In view of Lemma 3.4, this proves the lemma with $\alpha = \frac{2}{\text{HD}(\widehat{\Lambda}_0)}$ and with the hyperbolic set for f defined by $\Lambda = \bigcup_{i=0}^{M-1} f^i(\widehat{\Lambda})$. \square

3.3 Proof of the uniform scale lemma

Let $\varepsilon > 0$ be given. Let n_1 and $\eta_0 > 0$ be such that the conclusions of Proposition 3.1 hold with ε replaced by $\varepsilon/2$. Fix $\eta \in (0, \eta_0)$, and let α and Λ be given by Lemma 3.5 with η replaced by $\eta/6$. Since Λ is a hyperbolic set for f , there exist constants $C_0 > 0, \kappa > 0, \lambda > 1$ such that for every $x \in X$ and every integer $n \geq 1$ such that $\text{dist}(f^i(x), \Lambda) \leq 3\kappa$ for every $i \in \{0, 1, \dots, n-1\}$, $|Df^n(x)| \geq C_0 \lambda^n$ holds. It follows that there is a constant $C_1 > 0$ such that for every interval U intersecting Λ and satisfying $|U| \leq 3\kappa$, there is an integer $k \geq 0$ such that

$$k \leq C_1 \log(1/|U|), \quad 3\kappa \leq |f^k(U)| \leq 3\kappa \cdot \sup_X |Df|, \quad (8)$$

and such that f^k maps U diffeomorphically onto $f^k(U)$. Reduce κ if necessary, so that $\kappa \leq \eta/(3 \sup_X |Df|)$, and so that for every U and k as above we have in addition that the distortion of f^k on U is bounded by 2.

By Lemma 3.5, each ball of radius $\eta/6$ contains an α -safe point in Λ . From this and the compactness of X , we can find a finite subset F of $\Lambda \setminus E(\alpha)$ that is $(\eta/3)$ -dense in X . Let $n_0 \geq n_1$ be a sufficiently large integer so that F is disjoint from $E_{n_0}(\alpha)$,

$$n_0^{-\alpha} \leq \min \left\{ \frac{\eta}{6}, \frac{3}{2} \kappa \right\} \text{ and } \frac{n_0^{-\alpha}}{12\eta} \geq e^{-\frac{\varepsilon}{2}n_0}. \quad (9)$$

Now, let $n \geq n_0$ be an integer, and W a subinterval of X that satisfies $\eta \leq |f^n(W)| \leq 2\eta$. Since the finite set F is $(\eta/3)$ -dense, there is a point $x \in F$ whose distance to the mid point of $f^n(W)$ is at most $\eta/3$. Since $|f^n(W)| \geq \eta$ it follows that $B(x, \eta/6)$ is contained in $f^n(W)$. Together with the first inequality in (9) this implies that $U = B(x, n^{-\alpha})$ is contained in $f^n(W)$. Since by construction $x \notin E_{n_0}(\alpha)$, every pull-back of U by f^n is diffeomorphic. Take one pull-back of U by f^n contained in W and denote it by \widehat{J} .

Since $x \in \Lambda$ and $|U| = |B(x, n^{-\alpha})| \leq 3\kappa$ by the first inequality in (9), there is an integer $k \geq 0$ such that

$$k \leq C_1 \log(1/|U|) \leq C_1 \alpha \log n, \quad 3\kappa \leq |f^k(U)| \leq \eta$$

by (8), and such that f^k maps U diffeomorphically onto $f^k(U)$ with distortion bounded by 2. So, if we put $m = n + k$, then $n \leq m \leq n + C_1 \alpha \log n$ and f^m maps \widehat{J} diffeomorphically onto $f^m(\widehat{J})$. Denote by $J \subset W$ the pull-back by f^m of the interval with the same center as $f^m(\widehat{J})$ and whose length is equal to $\frac{1}{3}|f^m(\widehat{J})|$. By Proposition 3.1 with $n = m$ and $W = \widehat{J}$, the distortion of f^m on J is bounded by $e^{\varepsilon n}$. Note furthermore that

$$|f^m(J)| = \frac{1}{3}|f^m(\widehat{J})| = \frac{1}{3}|f^k(U)| \geq \kappa.$$

On the other hand, by Proposition 3.1 and the fact that the distortion of f^k on $U = f^n(\widehat{J})$ is bounded by 2, we have

$$\frac{n^{-\alpha}}{12\eta} \leq \frac{1}{6} \cdot \frac{|U|}{|f^n(W)|} \leq \frac{|f^n(J)|}{|f^n(W)|} \leq e^{\frac{\varepsilon}{2}n} \frac{|J|}{|W|}.$$

By the second inequality in (9) this implies $|J| \geq e^{-\varepsilon n}|W|$, and completes the proof of the lemma with $C = \alpha C_1$.

4 The large deviations upper bound

In this section we complete the proof of the large deviations upper bound in the Main Theorem. In Sect. 4.1 we construct certain horseshoes (Proposition 4.1) that are tailored to a given basic open set of $\mathcal{M}(f)$. The construction is based on the Uniform Scale Lemma in Sect. 3. In order to treat inflection critical points, initially we restrict ourselves to small intervals. In Sect. 4.2 we prove two intermediate estimates. The first is restricted to a small interval (Proposition 4.4), and the second is a global estimate (Proposition 4.6) obtained by spreading out the local estimate. In Sect. 4.3 we complete the large deviations upper bound.

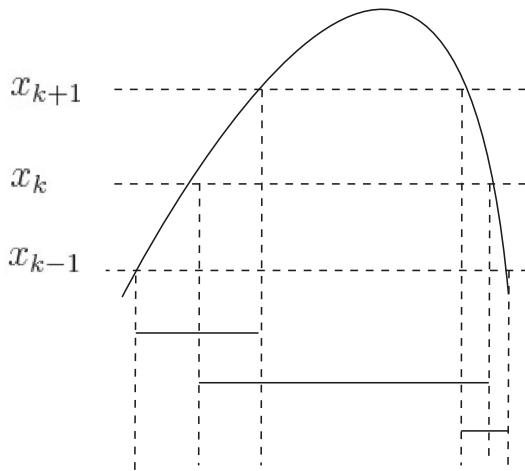
Positive constants we will be concerned with for the rest of this paper are $\varepsilon, \eta, \kappa, \rho$, chosen in this order. The purposes of them are as follows:

- ε is the error tolerance in the statement of Proposition 4.6;
- κ determines the scale of intervals given by the Uniform Scale Lemma;
- η determines the scale of the images of pull-backs of intervals;
- ρ determines the scale of horseshoes (see Proposition 4.1).

4.1 Horseshoe argument

Let $f : X \rightarrow X$ be a topologically exact continuous map. Let $n \geq 1$ be an integer and η in $(0, 1/2)$. Put $M = [1/\eta] + 1$ and note that $1/M < \eta < 3/(2M)$. Set $x_k = k/M$ for each $k \in \{1, 2, \dots, M-1\}$, and let $\mathcal{W}_n(x_k, \eta)$

Fig. 2 part of the graph of f^n and the partition of $\mathcal{P}_n(\eta)$. Every element of $\mathcal{P}_n(\eta)$ intersects no more than two other elements in their interiors



denote the collection of all pull-backs W of $B(x_k, \eta)$ by f^n that satisfy $x_k \in f^n(W)$. Note that elements of $\mathcal{W}_n(x_k, \eta)$ are pairwise disjoint. We now define

$$\mathcal{P}_n(\eta) = \bigcup_{k=1}^{M-1} \mathcal{W}_n(x_k, \eta).$$

It is easy to see that $\mathcal{P}_n(\eta)$ has the following properties:

- for every $x \in X$ there exists $W \in \mathcal{P}_n(\eta)$ such that $x \in W$;
- for every $W \in \mathcal{P}_n(\eta)$, we have $\eta \leq |f^n(W)| \leq 2\eta$;
- every element of $\mathcal{P}_n(\eta)$ can intersect at most two others on the boundary and two others in the interior. If $W_1, W_2 \in \mathcal{P}_n(\eta)$ and $\text{int}(W_1) \cap \text{int}(W_2) \neq \emptyset$, then for some $k \in \{2, \dots, M-1\}$,

$$\{W_1, W_2\} \subset \mathcal{W}_n(x_{k-1}, \eta) \cup \mathcal{W}_n(x_k, \eta) \cup \mathcal{W}_n(x_{k+1}, \eta).$$

The first two items follow from $f(X) = X$. The last one is immediate from the definitions, see Fig. 2.

Fix once and for all a point $x_0 \in \text{int } X$ such that $x_0 \notin \bigcup_{n=1}^{\infty} f^n(\text{Crit}(f))$.

Proposition 4.1 *Let $f: X \rightarrow X$ have Hölder continuous derivative and only non-flat critical points. Assume f is topologically exact. Then for every $\varepsilon_0 > 0$ there exist $\eta > 0$, $C > 0$ and $\rho > 0$ such that $B(x_0, 2\rho) \cap \partial X = \emptyset$, and the following holds. Let $l \geq 1$ be an integer, $\varphi_1, \dots, \varphi_l: X \rightarrow \mathbb{R}$ be continuous functions and let $\alpha_1, \dots, \alpha_l \in \mathbb{R}$. For each integer n define*

$$H_n = \left\{ x \in X : \text{for every } j \text{ in } \{1, \dots, l\} \text{ we have } \frac{1}{n} S_n \varphi_j(x) \geq \alpha_j \right\} \quad (10)$$

and

$$\mathcal{Q}_n = \{W \in \mathcal{P}_n(\eta) \text{ intersecting } H_n \cap B(x_0, \rho)\}.$$

Then, for each sufficiently large integer $n \geq 1$ such that \mathcal{Q}_n is nonempty, there exist an integer $q \geq n$ and pairwise disjoint diffeomorphic pull-backs L_1, \dots, L_t of $B(x_0, 2\rho)$ by f^q contained in $B(x_0, 2\rho)$ such that:

- (a) $n \leq q \leq n + C \log n$;
- (b) for each i in $\{1, \dots, t\}$ the distortion of f^q on L_i is bounded by $e^{\varepsilon_0 n}$, the interval L_i is contained in some $W \in \mathcal{Q}_n$, and $\sum_{W \in \mathcal{Q}_n} |W| \leq e^{\varepsilon_0 n} \sum_{i=1}^t |L_i|$;
- (c) for every $x \in \bigcup_{i=1}^t L_i$ and $j \in \{1, \dots, l\}$, we have $\frac{1}{q} S_q \varphi_j(x) > \alpha_j - \varepsilon_0$.

Proof Let $\varepsilon_0 > 0$. Since each φ_j ($j = 1, \dots, l$) is uniformly continuous, there exists $\varepsilon \in (0, \varepsilon_0)$ such that if $x, y \in X$ and $|x - y| \leq \varepsilon$ then $|\varphi_j(x) - \varphi_j(y)| \leq \varepsilon_0/2$. Let η_0, C and κ be the constants for which the conclusion of the Uniform Scale Lemma holds with ε replaced by $\varepsilon/4$. Fix $\eta \in (0, \eta_0)$ sufficiently small so that for every subinterval W of X and every integer $m \geq 1$ such that $|f^m(W)| \leq 2\eta$, we have for each $j \in \{0, \dots, m-1\}$ the estimate $|f^j(W)| \leq \varepsilon$ (Lemma 3.3). Recall that $N(\kappa) \geq 1$ is the smallest integer such that for every subinterval J of X with $|J| \geq \kappa$, $f^{N(\kappa)}(J) = X$, see Sect. 3. Let $\rho_0 > 0$ be sufficiently small such that $B(x_0, 2\rho_0) \cap \partial X = \emptyset$ and $B(x_0, 2\rho_0)$ is disjoint from $\bigcup_{i=1}^{N(\kappa)} f^i(\text{Crit}(f)) = \emptyset$. The last condition is indeed realized by our assumption $x_0 \notin \bigcup_{n=1}^{\infty} f^n(\text{Crit}(f))$, and it implies that each pull-back of $B(x_0, 2\rho_0)$ by $f^{N(\kappa)}$ is diffeomorphic. Let $\rho \in (0, \min\{\rho_0, \kappa\})$ be sufficiently small so that the distortion of $f^{N(\kappa)}$ on each pull-back of $B(x_0, 2\rho)$ by $f^{N(\kappa)}$ is bounded by 2.

Lemma 4.2 For every integer $n \geq N(\rho)$ and every $W \in \mathcal{P}_n(\eta)$ intersecting $B(x_0, \rho)$, we have $W \subset B(x_0, 2\rho)$.

Proof From the definition of $N(\rho)$ in Sect. 3, for every integer $n \geq N(\rho)$ and every pull-back $W \in \mathcal{P}_n(\eta)$, we have $|W| \leq \rho$. So $W \cap B(x_0, \rho) \neq \emptyset$ implies $W \subset B(x_0, 2\rho)$. \square

Let $n \geq \max\{n_0, N(\rho)\}$. By the Uniform Scale Lemma it is possible to choose for each $W \in \mathcal{Q}_n$ a closed subinterval $J_W \subset W$ and an integer $m_W \geq 1$ such that the following holds:

$$|J_W| \geq e^{-\frac{\varepsilon}{4}n} |W|, \quad n \leq m_W \leq n + C \log n, \quad |f^{m_W}(J_W)| \geq \kappa,$$

and f^{m_W} maps J_W diffeomorphically onto $f^{m_W}(J_W)$ with distortion bounded by $e^{\frac{\varepsilon}{4}n}$. Set

$$\mathcal{Q}_n(p) = \{W \in \mathcal{Q}_n : m_W = p\}.$$

Let p_0 denote a value of p that maximizes $\sum_{W \in \mathcal{Q}_n(p)} |W|$, so

$$\sum_{W \in \mathcal{Q}_n(p_0)} |W| \geq \frac{1}{1 + C \log n} \sum_{W \in \mathcal{Q}_n} |W|. \quad (11)$$

Set $q = p_0 + N(\kappa)$, and note that for every sufficiently large n item (a) holds with C replaced by $2C$. Since for each $W \in \mathcal{Q}_n(p_0)$ we have $|f^{p_0}(J_W)| \geq \kappa$, J_W contains at least one pull-back of $B(x_0, 2\rho)$ by f^q . Moreover, since the map $f^{p_0} : J_W \rightarrow f^{p_0}(J_W)$ is diffeomorphic, every pull-back of $B(x_0, 2\rho)$ by f^q that is contained in J_W is diffeomorphic. Pick one of these diffeomorphic pull-backs and denote it by L_W . Since by the Uniform Scale Lemma the distortion of f^{p_0} on J_W is bounded by $e^{\frac{\varepsilon}{4}n}$, and since by our choice of ρ the distortion of $f^{N(\kappa)} = f^{q-p_0}$ on $f^{p_0}(L_W)$ is bounded by 2, it follows that the distortion of f^q on L_W is bounded by $e^{\varepsilon n}$, provided that n is sufficiently large.

Lemma 4.3 *For every sufficiently large n and $W \in \mathcal{Q}_n(p_0)$, we have $|L_W| \geq e^{-\frac{3}{4}\varepsilon n} |W|$.*

Proof Since $|f^q(L_W)| = 4\rho$ and $q - n \leq 2C \log n$, we have

$$|f^n(L_W)| \geq |f^q(L_W)| \left(\sup_X |Df| \right)^{-(q-n)} \geq 4\rho \left(\sup_X |Df| \right)^{-2C \log n}.$$

Using $|f^n(J_W)| \leq |f^n(W)| \leq 2\eta$, and that the distortion of f^{p_0} on J_W is bounded by $e^{\frac{\varepsilon}{4}n}$, we also have

$$\frac{|L_W|}{|J_W|} \geq e^{-\frac{\varepsilon}{4}n} \frac{|f^n(L_W)|}{|f^n(J_W)|} \geq e^{-\frac{\varepsilon}{4}n} \frac{4\rho \left(\sup_X |Df| \right)^{-2C \log n}}{2\eta}.$$

Together with the inequality $|J_W| \geq e^{-\frac{\varepsilon}{4}n} |W|$, this completes the proof. \square

Any two elements of the collection of intervals $\{L_W : W \in \mathcal{Q}_n(p_0)\}$ are either disjoint or coincide with each other. Moreover, each of these intervals intersects at most five elements of $\{L_W : W \in \mathcal{Q}_n\}$. Let $\{L_i\}_{i=1}^t$ denote a collection of distinct elements of $\{L_W : W \in \mathcal{Q}_n(p_0)\}$ that maximizes $\sum_{i=1}^t |L_i|$.

Using (11) and Lemma 4.3, for every large integer $n \geq 1$ we have

$$\sum_{i=1}^t |L_i| \geq \frac{1}{5} \sum_{W \in \mathcal{Q}_n(p_0)} |L_W| \geq \frac{1}{5} e^{-\frac{3}{4}\varepsilon n} \sum_{W \in \mathcal{Q}_n(p_0)} |W| \geq e^{-\varepsilon n} \sum_{W \in \mathcal{Q}_n} |W|.$$

By Lemma 4.2, $L_i \subset B(x_0, 2\rho)$. Since $\varepsilon \in (0, \varepsilon_0)$ this completes the proof of item (b).

It is left to prove item (c). Since $L_W \subset J_W \subset W$ for every $W \in \mathcal{Q}_n(p_0)$, it suffices to prove the inequality for every $x \in \bigcup_{W \in \mathcal{Q}_n} W$. To ease notation, write φ, α for φ_j, α_j respectively. Let $W \in \mathcal{Q}_n$, choose a point $x \in W$ such that $S_n \varphi(x) \geq \alpha n$, and let $y \in W$. By our choice of η we have $|f^i(L_W)| \leq |f^i(W)| \leq \varepsilon$ for every $i \in \{0, \dots, n-1\}$, so

$$\frac{1}{n} |S_n \varphi(x) - S_n \varphi(y)| \leq \frac{\varepsilon_0}{2}.$$

Since

$$\begin{aligned} S_q \varphi(y) &= S_n \varphi(y) + S_{q-n} \varphi(f^n y) \\ &\geq S_n \varphi(x) - |S_n \varphi(x) - S_n \varphi(y)| - (q-n) \sup_X |\varphi| \end{aligned}$$

and $0 \leq q-n \leq 2C \log n$, for large n we have

$$\begin{aligned} \frac{1}{q} S_q \varphi(y) &\geq \frac{1}{q} S_n \varphi(x) - \frac{n \varepsilon_0}{2q} - \frac{q-n}{q} \sup_X |\varphi| \\ &\geq \frac{n}{q} \alpha - \frac{\varepsilon_0}{2} - \frac{2C \log n}{n} \sup_X |\varphi| \\ &> \alpha - \varepsilon_0. \end{aligned}$$

This completes the proof of item (c) and of the proposition. \square

4.2 Intermediate estimates

Using Proposition 4.1 we prove two propositions. The first one (Proposition 4.4) is a local estimate near the point x_0 chosen before Proposition 4.1. The second proposition (Proposition 4.6) is a global estimate that is obtained by using the topological exactness of f to spread out the local estimate.

Proposition 4.4 *Let $f : X \rightarrow X$ have Hölder continuous derivative and only non-flat critical points. Assume f is topologically exact. Then for every $\varepsilon_0 > 0$ there exists $\rho > 0$ such that the following holds. Let $l \geq 1$ be an integer,*

$\varphi_1, \dots, \varphi_l: X \rightarrow \mathbb{R}$ be continuous functions and let $\alpha_1, \dots, \alpha_l \in \mathbb{R}$. Then there exists an integer $n_0 \geq 1$ such that, if $n \geq n_0$ is an integer for which the set H_n defined by (10) is non-empty, then there exists $\mu \in \mathcal{M}(f)$ such that

$$\int \varphi_j d\mu > \alpha_j - \varepsilon_0 \text{ for every } j \in \{1, \dots, l\},$$

and

$$\frac{1}{n} \log |H_n \cap B(x_0, \rho)| \leq F(\mu) + \varepsilon_0.$$

The proof of this proposition is given after the following lemma. The next lemma will be proved along the standard line of the ergodic theory of uniformly hyperbolic systems.

Lemma 4.5 *Let $f: X \rightarrow X$ have continuous derivative and at most a finite number of critical points. Moreover, let B be a subinterval of X , $t, q \geq 1$ integers, and let L_1, \dots, L_t be pairwise disjoint diffeomorphic pull-backs of B by f^q contained in B . Finally, let $\Delta > 1$ be a constant such that for each i in $\{1, \dots, t\}$ the distortion of f^q on L_i is bounded by Δ . Then there exists $\widehat{\mu} \in \mathcal{M}(f^q)$ supported on $L_1 \cup \dots \cup L_t$, such that the measure $\mu = \frac{1}{q}(\widehat{\mu} + \dots + f_*^{q-1}\widehat{\mu})$ in $\mathcal{M}(f)$ satisfies*

$$qF(\mu) \geq \log \left(\frac{|L_1| + \dots + |L_t|}{\Delta |B|} \right).$$

Recall that for a continuous map $f: X \rightarrow X$, an integer $n \geq 1$ and $\varepsilon > 0$, a subset Y of X is (n, ε) -separated if for each distinct y and y' in Y there is j in $\{0, \dots, n-1\}$ such that $|f^j(y) - f^j(y')| \geq \varepsilon$.

Proof Let K be the maximal invariant set of f^q on $L_1 \cup \dots \cup L_t$, and fix a point y_0 in this set. Moreover, put

$$\varepsilon = \min\{\text{dist}(L_i, L_j) : i, j \in \{1, \dots, t\} \text{ distinct}\},$$

and note that for every integer $n \geq 1$ the set $(f^q|_K)^{-n}(y_0)$ is (n, ε) -separated for $f^q|_K$. From the definition of topological pressure in terms of (n, ε) -separated sets and the variational principle, this implies

$$\begin{aligned} & \sup_{\widehat{\nu} \in \mathcal{M}(f^q|_K)} \left(h_{f^q|_K}(\widehat{\nu}) - \int \log |Df^q| d\widehat{\nu} \right) \\ & \geq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \left(\sum_{x \in (f^q|_K)^{-n}(y_0)} |Df^{qn}(x)|^{-1} \right), \end{aligned}$$

where $\mathcal{M}(f^q|_K)$ denotes the set of $f^q|_K$ -invariant Borel probability measures and $h_{f^q|_K}(\widehat{\nu})$ denotes the entropy of $\widehat{\nu} \in \mathcal{M}(f^q|_K)$. See for example [25, Theorem 4.4.11] or [40, Theorems 3.3.2 and 3.4.1]. Using that for each i in $\{1, \dots, t\}$ the distortion of f^q on L_i is bounded by Δ , we have for every $n \geq 1$

$$\begin{aligned} \sum_{x \in (f^q|_K)^{-n}(y_0)} |Df^{qn}(x)|^{-1} &\geq \left(\inf_{y' \in K} \sum_{x' \in (f^q|_K)^{-1}(y')} |Df^q(x')|^{-1} \right)^n \\ &\geq \left(\frac{|L_1| + \dots + |L_t|}{\Delta|B|} \right)^n. \end{aligned}$$

We thus obtain

$$\sup_{\widehat{\nu} \in \mathcal{M}(f^q|_K)} \left(h_{f^q|_K}(\widehat{\nu}) - \int \log |Df^q| d\widehat{\nu} \right) \geq \log \left(\frac{|L_1| + \dots + |L_t|}{\Delta|B|} \right).$$

Since the measure-theoretic entropy of f^q is upper semi-continuous [34, Corollary 2], the supremum above is attained. Then the lemma follows from the fact that for each $\widehat{\nu}$ in $\mathcal{M}(f^q|_K)$, the measure $\nu = \frac{1}{q}(\widehat{\nu} + f_*\widehat{\nu} + \dots + f_*^{q-1}\widehat{\nu})$ is in $\mathcal{M}(f)$ and satisfies

$$h_{f^q|_K}(\widehat{\nu}) - \int \log |Df^q| d\widehat{\nu} = qF(\nu).$$

□

Proof of Proposition 4.4 Let $\varepsilon_0 > 0$. Take constants η, C, ρ , a positive integer q , and a collection of pairwise disjoint closed intervals L_1, \dots, L_t for which the conclusion of Proposition 4.1 holds with ε_0 replaced by $\varepsilon_0/2$. Since $H_n \cap B(x_0, \rho) \subset \bigcup_{W \in \mathcal{Q}_n} W$,

$$\log |H_n \cap B(x_0, \rho)| \leq \log \left(\sum_{W \in \mathcal{Q}_n} |W| \right).$$

Let $\mu \in \mathcal{M}(f^q)$ be as in Lemma 4.5 applied to $B = B(x_0, 2\rho)$, $\Delta = e^{\frac{\varepsilon_0}{2}n}$, and the pull-backs L_1, \dots, L_t of $B(x_0, 2\rho)$ by f^q . Proposition 4.1(c) yields $\int \varphi_j d\mu > \alpha_j - \varepsilon_0$ for every $j \in \{1, 2, \dots, t\}$. On the other hand, using $|B(x_0, \rho)| \leq 1$ and Proposition 4.1(b), for every large n we have

$$\log \left(\sum_{W \in \mathcal{Q}_n} |W| \right) \leq \log \left(\sum_{i=1}^t |L_i| \right) + \frac{\varepsilon_0}{2}n \leq qF(\mu) + \varepsilon_0 n.$$

Since $q \geq n$ and $F(\mu) \leq 0$ from Proposition 2.1, we have

$$\frac{1}{n} \log \left(\sum_{W \in \mathcal{Q}_n} |W| \right) \leq \frac{q}{n} F(\mu) + \varepsilon_0 \leq F(\mu) + \varepsilon_0.$$

This yields the desired inequality. \square

Proposition 4.6 *Let $f: X \rightarrow X$ have Hölder continuous derivative and only non-flat critical points. Assume f is topologically exact. Let $\varepsilon_0 > 0$, let $l \geq 1$ be an integer, let $\varphi_1, \dots, \varphi_l: X \rightarrow \mathbb{R}$ be continuous functions, and let $\alpha_1, \dots, \alpha_l \in \mathbb{R}$. Then*

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n} \log \left| \left\{ x \in X : \frac{1}{n} S_n \varphi_j(x) \geq \alpha_j \text{ for every } j \in \{1, \dots, l\} \right\} \right| \\ & \leq \sup \left\{ F(\mu) : \mu \in \mathcal{M}(f) \text{ and } \int \varphi_j d\mu > \alpha_j - \varepsilon_0 \text{ for every } j \in \{1, \dots, l\} \right\} \\ & \quad + \varepsilon_0. \end{aligned}$$

Remark 4.7 Since the Lyapunov exponent is not lower semi-continuous in general, it is not possible to let $\varepsilon_0 = 0$ in the inequality in Proposition 4.6.

Proof of Proposition 4.6 Let $\varepsilon_0 > 0$, $l \geq 1$, $\varphi_1, \dots, \varphi_l$, and $\alpha_1, \dots, \alpha_l$ be as in the statement of the proposition. Let $\rho > 0$ denote the constant for which the conclusion of Proposition 4.4 holds with ε_0 replaced by $\varepsilon_0/2$. Fix a large integer $M \geq 1$ with $f^M(B(x_0, \rho)) = X$. Since each φ_j is bounded, for sufficiently large n we have

$$\begin{aligned} & \left\{ x \in f^M(B(x_0, \rho)) : \frac{1}{n} S_n \varphi_j(x) \geq \alpha_j \text{ for every } j \in \{1, \dots, l\} \right\} \\ & \subset f^M \left\{ x \in B(x_0, \rho) : \frac{1}{n} S_n \varphi_j(x) \geq \alpha_j - \frac{\varepsilon_0}{2} \text{ for every } j \in \{1, \dots, l\} \right\}, \end{aligned}$$

and therefore

$$\begin{aligned} & \frac{1}{n} \log \left| \left\{ x \in X : \frac{1}{n} S_n \varphi_j(x) \geq \alpha_j \text{ for every } j \in \{1, \dots, l\} \right\} \right| \\ & \leq \frac{1}{n} \log \left[\left(\sup_X |Df| \right)^M \right. \\ & \quad \cdot \left. \left| \left\{ x \in B(x_0, \rho) : \frac{1}{n} S_n \varphi_j(x) \geq \alpha_j - \frac{\varepsilon_0}{2} \text{ for every } j \in \{1, \dots, l\} \right\} \right| \right] \end{aligned}$$

$$\leq \frac{1}{n} \log \left| \left\{ x \in B(x_0, \rho) : \frac{1}{n} S_n \varphi_j(x) \geq \alpha_j - \frac{\varepsilon_0}{2} \text{ for every } j \in \{1, \dots, l\} \right\} \right| + \frac{\varepsilon_0}{2}.$$

We use Proposition 4.4 with α_j replaced by $\alpha_j - \varepsilon_0/2$ for every $j \in \{1, \dots, l\}$. For each sufficiently large n there exists $\mu \in \mathcal{M}(f)$ such that $\int \varphi_j d\mu > \alpha_j - \varepsilon_0$ for every $j \in \{1, \dots, l\}$, and

$$\begin{aligned} & \frac{1}{n} \log \left| \left\{ x \in B(x_0, \rho) : \frac{1}{n} S_n \varphi_j(x) \geq \alpha_j - \frac{\varepsilon_0}{2} \text{ for every } j \in \{1, \dots, l\} \right\} \right| \\ & \leq F(\mu) + \frac{\varepsilon_0}{2}. \end{aligned}$$

Letting $n \rightarrow \infty$ we obtain the proposition. \square

4.3 End of the large deviations upper bound

We are in position to complete the large deviations upper bound in the Main Theorem.

Proof of the large deviations upper bound in the Main Theorem Let $f: X \rightarrow X$ have Hölder continuous derivative and only non-flat critical points, and assume it is topologically exact. Let \mathcal{K} be a closed subset of \mathcal{M} , and let \mathcal{G} be an arbitrary open set containing \mathcal{K} . Since \mathcal{K} is compact, one can choose a finite collection $\mathcal{C}_1, \dots, \mathcal{C}_r$ of closed sets such that $\mathcal{K} \subset \bigcup_{k=1}^r \mathcal{C}_k \subset \mathcal{G}$ and such that each of them has the form

$$\mathcal{C}_k = \left\{ \mu \in \mathcal{M} : \int \varphi_j d\mu \geq \alpha_j \text{ for every } j \in \{1, \dots, p\} \right\},$$

where $p \geq 1$ is an integer, each $\varphi_j: X \rightarrow \mathbb{R}$ is a continuous function and $\alpha_j \in \mathbb{R}$. For each $k \in \{1, 2, \dots, r\}$ and $\varepsilon_0 > 0$ define an open neighborhood $\mathcal{C}_k(\varepsilon_0)$ of \mathcal{C}_k by replacing $\int \varphi_j d\mu \geq \alpha_j$ in the definition of \mathcal{C}_k by $\int \varphi_j d\mu > \alpha_j - \varepsilon_0$. From Proposition 4.6, for every $\varepsilon_0 > 0$ and every $k \in \{1, 2, \dots, r\}$,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \left| \left\{ x \in X : \delta_x^n \in \mathcal{C}_k \right\} \right| \leq \sup_{\mathcal{C}_k(\varepsilon)} F + \varepsilon_0.$$

Since $\bigcup_{k=1}^r \mathcal{C}_k(\varepsilon_0) \subset \mathcal{G}$ for $\varepsilon_0 > 0$ small enough, using the previous inequality for each $k \in \{1, 2, \dots, r\}$ gives

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\{x \in X : \delta_x^n \in \mathcal{K}\}| &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \left| \left\{ x \in X : \delta_x^n \in \bigcup_{k=1}^r \mathcal{C}_k \right\} \right| \\ &\leq \max_{k \in \{1, 2, \dots, r\}} \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\{x \in X : \delta_x^n \in \mathcal{C}_k\}| \\ &\leq \max_{k \in \{1, 2, \dots, r\}} \sup_{\mathcal{C}_k(\varepsilon_0)} F + \varepsilon_0 \\ &\leq \sup_{\mathcal{G}} F + \varepsilon_0. \end{aligned}$$

Letting $\varepsilon_0 \rightarrow 0$ we obtain

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\{x \in X : \delta_x^n \in \mathcal{K}\}| \leq \sup_{\mathcal{G}} F.$$

Since \mathcal{G} is an arbitrary open set containing \mathcal{K} , it follows that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\{x \in X : \delta_x^n \in \mathcal{K}\}| \leq \inf_{\mathcal{G} \supset \mathcal{K}} \sup_{\mathcal{G}} F = \inf_{\mathcal{G} \supset \mathcal{K}} \sup_{\mathcal{G}} (-I) = -\inf_{\mathcal{K}} I.$$

The last equality is due to the upper semi-continuity of $-I$. \square

Acknowledgements We would like to thank Michał Misiurewicz for his help with references, Bing Gao, Gerhard Keller and Masato Tsujii for fruitful discussions, and the anonymous referees for their healthy criticism that helped us improve the exposition in the introduction. The first-named author is partially supported by the Grant-in-Aid for Scientific Research (C) of the JSPS 16K05179. The second-named author is partially supported by FONDECYT Grant 1141091 and NSF Grant DMS-1700291. The last-named author is partially supported by the Grant-in-Aid for Young Scientists (A) of the JSPS 15H05435 and the Grant-in-Aid for Scientific Research (B) of the JSPS 16KT0021.

Appendix Rate functions for Hofbauer-Keller maps

Let $f_a : X \rightarrow X$ ($0 < a \leq 4$) be the quadratic map $f_a(x) = ax(1-x)$. Let $c = 1/2$ and put $X_a = [f_a^2(c), f_a(c)]$. Notice that $f_a(X_a) = X_a$. Denote by \mathcal{M}_a the space of Borel probability measures on X_a endowed with the weak* topology, and by $\mathcal{M}_a(f_a)$ the set of elements of \mathcal{M}_a which are $f_a|_{X_a}$ -invariant.

By [19, 31], for Lebesgue almost every $a \in (0, 4]$ there exists a unique physical measure of f_a . Based on the kneading theory, Hofbauer & Keller [21, 22] constructed various examples of quadratic maps with unexpected properties. One of them is the following.

Theorem A.1 ([22], Propositions 1 and 2) *There is a uncountable set $A \subset (0, 4)$ such that if $a \in A$ then f_a is non-renormalizable and there are sequences $\{n_i\}_i$, $\{m_i\}_i$ of positive integers with $n_i < m_i < n_{i+1}$ for each i such that the following holds:*

- (a) $\left| \int \varphi d\delta_x^{n_i} - \int \varphi d\delta_c^{m_i} \right| \rightarrow 0$ ($i \rightarrow \infty$) for Lebesgue almost every $x \in X_a$ and each continuous $\varphi: X_a \rightarrow \mathbb{R}$;
- (b) if $z \in X_a$ and $p \geq 1$ are such that $f_a^p(z) = z$, then δ_z^p is a weak*-accumulation point of the sequence $\{\delta_c^{m_i}\}_{i \geq 1}$.

In particular, if $a \in A$ then there is no physical measure of f_a . Hence, the law of large numbers does not hold for the Birkhoff sum $\varphi + \varphi \circ f_a + \dots + \varphi \circ f_a^{n-1}$ of a continuous function $\varphi: X_a \rightarrow \mathbb{R}$. Nevertheless, $f_a|_{X_a}$ satisfies the hypotheses of the Main Theorem and hence the LDP holds. The rate function is identically zero on its effective domain.

Theorem A.2 *Let A be the set as in Theorem A.1. If $a \in A$ then the large deviations rate function of $f_a|_{X_a}$ is identically zero on $\mathcal{M}_a(f_a)$.*

Proof Let $a \in A$ and $\mu \in \mathcal{M}_a(f_a)$. Let \mathcal{U} be an arbitrary open set containing μ . Take $l \geq 1$, continuous functions $\varphi_1, \dots, \varphi_l: X_a \rightarrow \mathbb{R}$, $\varepsilon > 0$ such that $\mu \in \mathcal{C} \subset \mathcal{U}$, where

$$\mathcal{C} = \left\{ \nu \in \mathcal{M}_a: \left| \int \varphi_j d\nu - \int \varphi_j d\mu \right| \leq \varepsilon \text{ for every } j \in \{1, \dots, l\} \right\}.$$

Since f_a is non-renormalizable, its restriction to X_a is topologically exact and has the specification property. Hence μ is weak*-approximated by another supported on a periodic orbit [46, Theorem 1] and there exist $z \in X_a$ and $p \geq 1$ such that $f_a^p(z) = z$ and

$$\left| \int \varphi_j d\delta_z^p - \int \varphi_j d\mu \right| \leq \frac{\varepsilon}{3} \text{ for every } j \in \{1, \dots, l\}.$$

From Theorem A.1 there are increasing sequences $\{n_i\}_i$, $\{m_i\}_i$ of positive integers for which the following holds:

$$\begin{aligned} \left| \left\{ x \in X_a: \left| \int \varphi_j d\delta_x^{n_i} - \int \varphi_j d\delta_c^{m_i} \right| \leq \frac{\varepsilon}{3} \text{ for every } j \in \{1, \dots, l\} \right\} \right| &\geq \frac{1}{2}; \\ \left| \int \varphi_j d\delta_c^{m_i} - \int \varphi_j d\delta_z^p \right| &\leq \frac{\varepsilon}{3} \text{ for every } j \in \{1, \dots, l\}. \end{aligned}$$

Combining these three inequalities yields

$$\begin{aligned} \frac{1}{2} &\leq \left| \left\{ x \in X_a : \left| \int \varphi_j d\delta_x^{n_i} - \int \varphi_j d\mu \right| \leq \varepsilon \text{ for every } j \in \{1, \dots, l\} \right\} \right| \\ &\leq |X_a| \leq 1. \end{aligned} \quad (12)$$

Denote by I_a the large deviations rate function of $f_a|_{X_a}$. Then

$$0 \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\{x \in X_a : \delta_x^n \in \mathcal{C}\}| \leq -\inf_{\mathcal{C}} I_a \leq -\inf_{\mathcal{U}} I_a \leq 0.$$

The first inequality is from (12) and the second from the Main Theorem. Hence $\inf_{\mathcal{U}} I_a = 0$. Since \mathcal{U} is an arbitrary open set containing μ and I_a is lower semi-continuous, $I_a(\mu) = 0$. \square

References

1. Avila, A., Lyubich, M., de Melo, W.: Regular or stochastic dynamics in real analytic families of unimodal maps. *Invent. Math.* **154**(3), 451–550 (2003)
2. Benedicks, M., Carleson, L.: On iterations of $1 - ax^2$ on $(-1, 1)$. *Ann. Math.* **122**(1), 1–25 (1985)
3. Blokh, A., Oversteegen, L.: Backward stability for polynomial maps with locally connected Julia sets. *Trans. Am. Math. Soc.* **356**(1), 119–133 (2004)
4. Bowen, R.: Equilibrium States and The Ergodic Theory of Anosov Diffeomorphisms. Lecture Notes in Mathematics, vol. 470. Springer, Berlin (1975)
5. Bruin, H., Keller, G.: Equilibrium states for S -unimodal maps. *Ergod. Theory Dyn. Syst.* **18**(4), 765–789 (1998)
6. Cai, H., Li, S.: Distortion of interval maps and applications. *Nonlinearity* **22**(10), 2353–2363 (2009)
7. Chung, Y.M.: Large deviations on Markov towers. *Nonlinearity* **24**(4), 1229–1252 (2011)
8. Chung, Y.M., Takahasi, H.: Large deviation principle for Benedicks–Carleson quadratic maps. *Commun. Math. Phys.* **315**(3), 803–826 (2012)
9. Chung, Y.M., Takahasi, H.: Multifractal formalism for Benedicks–Carleson quadratic maps. *Ergod. Theory Dyn. Syst.* **34**(4), 1116–1141 (2014)
10. Chung, Y.M., Takahasi, H.: Large deviation principle for S -unimodal maps with flat critical point. [arXiv:1708.03695v2](https://arxiv.org/abs/1708.03695v2) (2017)
11. Collet, P., Eckmann, J.-P.: Positive Liapunov exponents and absolute continuity for maps of the interval. *Ergod. Theory Dyn. Syst.* **3**(1), 13–46 (1983)
12. Comman, H., Rivera-Letelier, J.: Large deviation principles for non-uniformly hyperbolic rational maps. *Ergod. Theory Dyn. Syst.* **31**(2), 321–349 (2011)
13. de Melo, W., van Strien, S.: One-Dimensional Dynamics. *Ergebnisse der Mathematik und ihrer Grenzgebiete* (3), vol. 25. Springer, Berlin (1993)
14. Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications. 2nd edn. *Applications of Mathematics* (New York), vol. 38. Springer, New York (1998)
15. Denker, M.: Probability Theory for Rational Maps. In: Probability theory and mathematical statistics (St. Petersburg, 1993), pp. 29–40. Gordon and Breach, Amsterdam (1996)
16. Dobbs, N.: On cusps and flat tops. *Ann. Inst. Fourier (Grenoble)* **64**(2), 571–605 (2014)

17. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time. I. II. *Commun. Pure Appl. Math.* **28**, 1–47 (1975); *ibid.* **28**, 279–301 (1975)
18. Ellis, R.S.: Entropy, Large Deviations, and Statistical Mechanics, Grundlehren der Mathematischen Wissenschaften, vol. 271. Springer, New York (1985)
19. Graczyk, J., Świątek, G.: Generic hyperbolicity in the logistic family. *Ann. Math.* **146**(1), 1–52 (1997)
20. Grigull, J.: Große Abweichungen und Fluktuationen für Gleichgewichtsmaße rationaler Abbildungen. Ph.D. thesis (1993)
21. Hofbauer, F., Keller, G.: Quadratic maps without asymptotic measure. *Commun. Math. Phys.* **127**(2), 319–337 (1990)
22. Hofbauer, F., Keller, G.: Quadratic Maps with Maximal Oscillation. In: Algorithms, fractals, and dynamics (Okayama/Kyoto, 1992), pp. 89–94. Plenum, New York (1995)
23. Jakobson, M.V.: Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. *Commun. Math. Phys.* **81**(1), 39–88 (1981)
24. Katok, A., Hasselblatt, B.: Introduction to the Modern Theory of Dynamical Systems. With a supplementary chapter by Katok and Leonardo Mendoza. Encyclopedia of Mathematics and its Applications, vol. 54. Cambridge University Press, Cambridge (1995)
25. Keller, G.: Equilibrium States in Ergodic Theory. London Mathematical Society Student Texts, vol. 42. Cambridge University Press, Cambridge (1998)
26. Keller, G., Nowicki, T.: Spectral theory, zeta functions and the distribution of periodic points for Collet-Eckmann maps. *Commun. Math. Phys.* **149**(1), 31–69 (1992)
27. Keller, G., Nowicki, T.: Fibonacci maps re(al)visited. *Ergod. Theory Dyn. Syst.* **15**(1), 99–120 (1995)
28. Kifer, Y.: Large deviations in dynamical systems and stochastic processes. *Trans. Am. Math. Soc.* **321**(2), 505–524 (1990)
29. Levin, G.: On backward stability of holomorphic dynamical systems. *Fund. Math.* **158**(2), 97–107 (1998)
30. Li, H.: Large deviation principles of one-dimensional maps for Hölder continuous potentials. *Ergod. Theory Dyn. Syst.* **36**(1), 127–141 (2016)
31. Lyubich, M.: Almost every real quadratic map is either regular or stochastic. *Ann. Math.* **156**(1), 1–78 (2002)
32. Lyubich, M., Milnor, J.: The Fibonacci unimodal map. *J. Am. Math. Soc.* **6**(2), 425–457 (1993)
33. Melbourne, I., Nicol, M.: Large deviations for nonuniformly hyperbolic systems. *Trans. Am. Math. Soc.* **360**(12), 6661–6676 (2008)
34. Misiurewicz, M., Szlenk, W.: Entropy of piecewise monotone mappings. *Studia Math.* **67**(1), 45–63 (1980)
35. Orey, S., Pelikan, S.: Deviations of trajectory averages and the defect in Pesin's formula for Anosov diffeomorphisms. *Trans. Am. Math. Soc.* **315**(2), 741–753 (1989)
36. Pollicott, M., Sharp, R., Yuri, M.: Large deviations for maps with indifferent fixed points. *Nonlinearity* **11**(4), 1173–1184 (1998)
37. Przytycki, F., Rivera-Letelier, J.: Nice inducing schemes and the thermodynamics of rational maps. *Commun. Math. Phys.* **301**(3), 661–707 (2011)
38. Przytycki, F., Rivera-Letelier, J.: Geometric pressure for multimodal maps of the interval. *Mem. Am. Math. Soc.* **259**(1246), v+81 (2019)
39. Przytycki, F., Rivera-Letelier, J., Smirnov, S.: Equivalence and topological invariance of conditions for non-uniform hyperbolicity in the iteration of rational maps. *Invent. Math.* **151**(1), 29–63 (2003)
40. Przytycki, F., Urbański, M.: Conformal Fractals: Ergodic Theory Methods. London Mathematical Society Lecture Note Series, vol. 371. Cambridge University Press, Cambridge (2010)

41. Rey-Bellet, L., Young, L.-S.: Large deviations in non-uniformly hyperbolic dynamical systems. *Ergod. Theory Dyn. Syst.* **28**(2), 587–612 (2008)
42. Rivera-Letelier, J.: Asymptotic expansion of smooth interval maps. [arXiv:1204.3071v2](https://arxiv.org/abs/1204.3071v2) (2012)
43. Ruelle, D.: A measure associated with axiom-A attractors. *Am. J. Math.* **98**(3), 619–654 (1976)
44. Ruelle, D.: An inequality for the entropy of differentiable maps. *Bol. Soc. Brasil. Mat.* **9**(1), 83–87 (1978)
45. Ruette, S.: Chaos on the Interval. University Lecture Series, vol. 67. American Mathematical Society, Providence, RI (2017)
46. Sigmund, K.: On dynamical systems with the specification property. *Trans. Am. Math. Soc.* **190**, 285–299 (1974)
47. Sinaĭ, J.G.: Gibbs measures in ergodic theory. *Uspehi Mat. Nauk* **27**(4(166)), 21–64 (1972)
48. Takahashi, Y.: Entropy Functional (free energy) for Dynamical Systems and their Random Perturbations. In: Stochastic analysis (Katata/Kyoto, 1982), North-Holland Math. Library, vol. 32, pp. 437–467. North-Holland, Amsterdam (1984)
49. Takahashi, Y.: Asymptotic Behaviours of Measures of Small Tubes: Entropy, Liapunov’s Exponent and Large Deviation. In: Dynamical systems and applications (Kyoto, 1987), World Sci. Adv. Ser. Dynam. Systems, vol. 5, pp. 1–21. World Sci. Publishing, Singapore (1987)
50. Young, L.-S.: Large deviations in dynamical systems. *Trans. Am. Math. Soc.* **318**(2), 525–543 (1990)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.