Large and Moderate Deviations for Slowly Mixing Dynamical Systems

Ian Melbourne *

7 June, 2008

Abstract

We obtain results on large deviations for a large class of nonuniformly hyperbolic dynamical systems with polynomial decay of correlations $1/n^\beta$, $\beta > 0$. This includes systems modelled by Young towers with polynomial tails, extending recent work of M. Nicol and the author which assumed $\beta > 1$. As a byproduct of the proof, we obtain slightly stronger results even when $\beta > 1$. The results are sharp in the sense that there exist examples (such as Pomeau-Manneville intermittency maps) for which the obtained rates are best possible. In addition, we obtain results on moderate deviations.

1 Introduction

Let $T : X \to X$ be a dynamical system with ergodic invariant probability measure μ. Suppose that $\phi : X \to \mathbb{R}$ is an L^1 observable with mean $\bar{\phi} = \int_X \phi \, d\mu$. Let $\phi_N = \sum_{j=0}^{N-1} \phi \circ T^j$. Birkhoff's pointwise ergodic theorem guarantees that $\lim_{N\to\infty} \frac{1}{N}\phi_N = \bar{\phi}$ almost everywhere, and this implies the weak law of large numbers $\lim_{N\to\infty} \mu(|\frac{1}{N}\phi_N - \bar{\phi}| > \epsilon) = 0$ for all $\epsilon > 0$. Large deviations theory concerns the rate of decay in the weak law of large numbers, with an emphasis on exponential convergence tied to thermodynamic formalism [6, 7].

For uniformly hyperbolic dynamical systems there is a good theory of thermodynamic formalism and results on large deviations were obtained by [11, 13, 17, 23, 24]. A general class of one-dimensional maps was considered by [10].

For nonuniformly hyperbolic systems, recent progress by [1, 2] yields strong results when it is known that there is a unique equilibrium measure. A different approach by [15] exploits quasicompactness (following [8]) and yields exponential large deviation results for Hölder observables of nonuniformly hyperbolic systems modelled by Young

*Department of Mathematics, University of Surrey, Guildford GU2 7XH, UK, ism@math.uh.edu

towers with exponential tails [25]. Subsequently, [21] obtained similar results by independently following the same approach as [8, 15].

The work of Melbourne & Nicol [15] also addressed subexponential decay rates for large deviations. Noting [15, Formula (3.2)], their main result in this direction can be expressed as follows.

Theorem 1.1 ([15]) *Let $\beta > 1$. Let $\phi \in L^\infty(X)$ and suppose that*

$$|\int_X \phi\,\psi \circ T^n\,d\mu - \int_X \phi\,d\mu \int_X \psi\,d\mu| \le C_\phi \|\psi\|_\infty n^{-\beta},$$

for all $\psi \in L^\infty(X)$, $n \ge 1$. Then for any $\delta > 0$,

$$\mu(|\tfrac{1}{N}\phi_N - \bar{\phi}| > \epsilon) \le C_{\phi,\epsilon,\delta}\, N^{-(\beta-\delta)}, \;\; \textit{for all } N \ge 1.$$

In other words, if ϕ has polynomial decay of correlations against all L^∞ test functions, then ϕ has large deviations decaying at essentially the same rate.

Young [26] considered a class of nonuniformly expanding dynamical systems (modelled by Young towers with polynomial tails) for which the hypothesis of Theorem 1.1 holds for Hölder observables ϕ. Hence such observables have polynomial large deviations. Moreover, [15] gave examples where there is a nonempty open set of Hölder observables such that $N^{-\beta}$ is a lower bound for sufficiently large N, so the result in [15] is close to optimal.

The hypothesis of Theorem 1.1 cannot hold for invertible maps. Nevertheless there is a class of nonuniformly hyperbolic diffeomorphisms modelled by Young towers with polynomial tails, and standard approximation arguments reduce both decay of correlations (against Hölder test functions) and large deviations to the noninvertible situation. (We refer to [15] for examples where these results apply.)

The method in [15] requires $\beta > 1$ (summable decay of correlations). In this paper, we obtain large deviations for all $\beta > 0$. Unexpectedly, the more general argument presented here gives slightly better results even for $\beta > 1$.

Theorem 1.2 (Large deviations) *Let $\beta > 0$. Let $\phi \in L^\infty(X)$ and suppose that*

$$|\int_X \phi\,\psi \circ T^n\,d\mu - \int_X \phi\,d\mu \int_X \psi\,d\mu| \le C_\phi \|\psi\|_\infty n^{-\beta},$$

for all $\psi \in L^\infty(X)$, $n \ge 1$. Then for any $\epsilon > 0$,

$$\mu(|\tfrac{1}{N}\phi_N - \bar{\phi}| > \epsilon) \le C_{\phi,\epsilon} N^{-\beta}, \;\; \textit{for all } N \ge 1.$$

We also obtain results on moderate deviations.

Theorem 1.3 (Moderate deviations) *Assume the hypotheses of Theorem 1.2 and let $\tau \in (\frac{1}{2}, 1]$. Then*

$$\mu\Big(\big|\tfrac{1}{N^\tau}(\phi_N - N\bar{\phi})\big| > \epsilon\Big) \le \begin{cases} C_{\phi,\epsilon}(\ln N)^\beta N^{-\beta(2\tau-1)}, & \beta \ge 1 \\ C_{\phi,\epsilon} N^{-(\beta-2(1-\tau))}, & \beta < 1 \end{cases}$$

for all $N \ge 1$.

Remark 1.4 As in [15], Theorems 1.2 and 1.3 apply immediately to nonuniformly expanding and nonuniformly hyperbolic systems modelled by Young towers with polynomial tails.

Remark 1.5 We have suppressed the dependence of the constants $C_{\phi,\epsilon}$ (and $C_{\phi,\epsilon,\delta}$) on β and τ. As was the case in [15], the constants $C_{\phi,\epsilon}$ are continuous functions of C_ϕ and ϵ (explicit formulas are given in Section 2). In particular, the ϵ dependence has the form $C_{\beta/q}\,\epsilon^{-2q}$ where $q > \max\{1, \beta\}$ in Theorem 1.2 and $q = \max\{1, \beta\}$ in Theorem 1.3. (But note Remark 2.2.) In the applications to nonuniformly hyperbolic systems, C_ϕ is a scalar multiple of the Hölder constant of ϕ.

Example 1.6 Intermittency (Pomeau-Manneville) maps $T : [0, 1] \to [0, 1]$ with an indifferent fixed point at 0 have been studied by various authors including [9, 12, 20, 26]. These maps are given by

$$Tx = \begin{cases} x(1 + 2^\alpha x^\alpha), & 0 \le x < \frac{1}{2} \\ 2x - 1, & \frac{1}{2} \le x < 1 \end{cases}$$

for $\alpha \in (0, 1)$. There is a unique ergodic invariant probability measure μ equivalent to Lebesgue and the hypotheses of Theorems 1.2 and 1.3 are satisfied with $\beta = \frac{1}{\alpha} - 1$. Hence we obtain large and moderate deviations for all $\alpha \in (0, 1)$. These results are completely new for $\alpha \in [\frac{1}{2}, 1)$ (the regime where the central limit theorem fails) and slightly improve existing results [15] for $\alpha \in (0, \frac{1}{2})$.

By Melbourne & Nicol [15, Proposition 3.3], there is a nonempty open set of Hölder observables ϕ for which $N^{-\beta}$ is a lower bound for large deviations for N sufficiently large. For these observables, we have

$$\lim_{N\to\infty} \frac{\log \mu(|\frac{1}{N}\phi_N - \bar{\phi}| > \epsilon)}{\log N} = -\beta.$$

Moreover, it follows that there is an open and dense set of Hölder observables ϕ for which $N^{-\beta}$ is a lower bound for infinitely many values of N. (This is proved in the appendix, slightly improving [15, Theorem 3.5].) For these observables, we have

$$\limsup_{N\to\infty} \frac{\log \mu(|\frac{1}{N}\phi_N - \bar{\phi}| > \epsilon)}{\log N} = -\beta.$$

Example 1.7 Planar billiards are an important class of examples in mathematical physics, and provide a number of situations where the hypotheses of Theorems 1.2 and 1.3 are satisfied with $\beta = 1$: Bunimovich stadia [5, 14], dispersing billiards with cusps [3, 5], and certain classes of semi-dispersing billiards [4, 5]. We are not aware of previous results on large or moderate deviations for any of these examples.

Remark 1.8 The moderate deviation results give estimates for $\tau > \frac{1}{2}$ if $\beta \ge 1$, and for $\tau > 1 - \beta/2$ if $0 < \beta < 1$. It is not likely that these estimates are optimal at least for $\beta < 1$.

Remark 1.9 Independently, Pollicott & Sharp [18] have obtained similar results for intermittency maps, building upon previous work of [19]. For $\beta \le 1$ in Example 1.6, they too obtain the optimal rate $1/N^\beta$. For $\beta > 1$, their upper bound improves upon [15] but is weaker than our rate (which is optimal).

Pollicott & Sharp [18] also consider *Level II* results for measures, in addition to *Level I* results for functions of the type considered in this paper. Indeed, the argument of [18, Theorem 3] to deduce *Level II* from *Level I* goes through here without change. Hence we obtain the following result which applies to all nonuniformly hyperbolic dynamical systems modelled by Young towers (including Examples 1.6 and 1.7), and hence generalises and strengthens [18, Theorem 3].

Theorem 1.10 (Level II) *Suppose that X is a compact metric space and that the hypothesis (or conclusion) of Theorem 1.2 holds for all Hölder continuous observables ϕ of some fixed Hölder exponent. Let $\mathcal{M}$ be the set of Borel probability measures on X with the weak* topology.*

If $\mathcal{K} \subset \mathcal{M}$ is a compact subset with $\mu \notin \mathcal{K}$, then

$$\mu\Big\{x \in X : \frac{1}{N}\sum_{j=0}^{N-1} \delta_{T^j x} \in \mathcal{K}\Big\} \le C_{\mathcal{K}} N^{-\beta}, \quad \text{for all } N \ge 1.$$

∎

2 Proof of the main results

Let $T : X \to X$ be a measurable transformation with Koopman operator $U : L^2(X) \to L^2(X)$ given by $U\phi = \phi \circ T$ and transfer operator $P = U^* : L^2(X) \to L^2(X)$.

Lemma 2.1 *Let $\beta > 0$ and $q \ge 1$. Let $\phi \in L^\infty$ with $\int_X \phi\, d\mu = 0$. Suppose that $|\int_X \phi\, \psi \circ T^n\, d\mu| \le C_\phi \|\psi\|_\infty n^{-\beta}$ for all $\psi \in L^\infty(X)$, $n \ge 1$.*

Then $\|\phi_N\|_{2q}^{2q} \le C_{\beta,q} C_\phi^q \|\phi\|_\infty^{2q-1} g(N)$ for N sufficiently large, where

$$g(N) = \begin{cases} N^{2q-\beta}, & q > \beta \\ N^q (\ln N)^q, & q = \beta \end{cases}.$$

Proof Since $P = U^*$ we have $|\int_X P^n\phi\, \psi\, d\mu| \le C_\phi \|\psi\|_\infty n^{-\beta}$ for all $\psi \in L^\infty(X)$, $n \ge 1$. Taking $\psi = \operatorname{sgn} P^n\phi$, we obtain $\|P^n\phi\|_1 \le C_\phi n^{-\beta}$. Hence $\int_X |P^n\phi|^q\, d\mu \le \|P^n\phi\|_\infty^{q-1} \int_X |P^n\phi|\, d\mu \le \|\phi\|_\infty^{q-1} C_\phi n^{-\beta}$. It follows that $\|P^n\phi\|_q \le C'_\phi n^{-\beta/q}$ where $C'_\phi = C_\phi^{1/q} \|\phi\|_\infty^{1-1/q}$.

Define

$$\chi^k = \sum_{j=1}^{k} P^j \phi, \qquad \psi^k = \phi + \chi^k - \chi^k \circ T - P^k\phi. \tag{2.1}$$

Then $\|\chi^k\|_q \le C'_\phi g_0(k)$ and $\|\psi^k\|_q \le 2C'_\phi g_0(k)$ for k sufficiently large where

$$g_0(k) = \begin{cases} 2(1-\beta/q)^{-1}k^{1-\beta/q}, & q > \beta \\ 2\ln k, & q = \beta \end{cases}.$$

We compute that $P\psi^k = 0$ so that $\{\psi^k \circ T^j;\ j = 0, 1, 2, \ldots\}$ is a sequence of reverse martingale differences. As usual, we may pass to the natural extension, and so may suppose without loss that $\{\psi^k \circ T^j;\ j = 0, 1, 2, \ldots\}$ is a sequence of martingale differences with respect to a filtration $\{\mathcal{F}_j;\ j = 0, 1, 2, \ldots\}$.

Rio's inequality [22, Theorem 2.5] (or [16, Proposition 7]) guarantees that $\|\phi_N\|_{2q}^{2q} \le \big(4q\sum_{i=1}^N b_{i,N}\big)^q$, where

$$b_{i,N} = \max_{i\le u\le N} \|\phi \circ T^i \sum_{\ell=i}^u E(\phi\circ T^\ell|\mathcal{F}_i)\|_q \le \|\phi\|_\infty \max_{i\le u\le N} \|\sum_{\ell=i}^u E(\phi\circ T^\ell|\mathcal{F}_i)\|_q.$$

It follows from the decomposition (2.1) that

$$\sum_{\ell=i}^u E(\phi\circ T^\ell|\mathcal{F}_i) = \psi^k\circ T^i + E(\chi^k\circ T^{u+1}|\mathcal{F}_i) - E(\chi^k\circ T^i|\mathcal{F}_i) + \sum_{\ell=i}^u E(P^k\phi\circ T^\ell|\mathcal{F}_i),$$

so that

$$\|\sum_{\ell=i}^u E(\phi\circ T^\ell|\mathcal{F}_i)\|_q \le \|\psi^k\|_q + 2\|\chi^k\|_q + N\|P^k\phi\|_q \le C'_\phi(4g_0(k) + Nk^{-\beta/q}).$$

Hence $b_{i,N} \le C''_\phi(4g_0(k) + Nk^{-\beta/q})$, where $C''_\phi = \|\phi\|_\infty^{2-1/q}C_\phi^{1/q}$. Taking $k = N$ yields the required (optimal) estimates. ∎

Proof of Theorems 1.2 and 1.3 Without loss, we may suppose that $\bar\phi = 0$. Hence we are in the situation of Lemma 2.1. By Markov's inequality, $\mu(|\frac{1}{N^\tau}\phi_N| > \epsilon) \le \epsilon^{-2q}N^{-2\tau q}\|\phi_N\|_{2q}^{2q}$. We choose $q \ge 1$ in accordance with Lemma 2.1. For the large deviation estimate, take $q > \max\{1, \beta\}$. For the moderate deviation estimate, set $q = \max\{1, \beta\}$. ∎

Remark 2.2 Note that the moderate deviation estimate in Theorem 1.3 gives an alternative estimate for large deviations (setting $\tau = 1$) which is stronger in ϵ but weaker in N. Taking $q < \beta$ (and $k = \infty$) recovers the estimate $O(\epsilon^{-2q}N^{-q})$ obtained in [15] which is even stronger in ϵ but even weaker in N.

A Lower bounds

The following result is proved in [15] for certain maps $T : X \to X$ (such as those in Example 1.6) that are particularly well-modelled by Young towers, see [15, Remark 3.6]. In such situations the following results hold.

Proposition A.1 ([15, Proposition 3.3]) *There exists $c_0 > 0$ with the property that for any $\delta > 0$ there exists $N_0 \geq 1$ and ϕ Hölder with mean zero and $\|\phi\| < \delta$ such that $\mu(|\frac{1}{N}\phi_N| > \delta/2) \geq c_0/N^\beta$ for all $N \geq N_0$. Moreover, this estimate holds for all nearby Hölder observables ϕ.*

Proof Assuming that we are in the situation of [15, Remark 3.6], it suffices to work on the Young tower itself. We assume familiarity with the notation of Young towers [26] as used in [15]. In particular, let Δ be a tower with base Y, invariant measure μ_Δ and return time function $R : Y \to \mathbb{Z}^+$. By assumption, $\mu_\Delta(y \in Y : R(y) > n\} \sim N^{-(\beta+1)}$. For $N \geq 1$, define $D_N = \{(y,\ell) \in \Delta : R(y) \geq N\}$, and $E_N = \{(y,\ell) \in D_N : \ell < R(y) - N - 1\}$. Then $\mu_\Delta(D_N) \sim N^{-\beta}$ and $\mu_\Delta(E_N) \sim N^{-\beta}$. We choose $c_0 > 0$ so that $\mu_\Delta(E_N) \geq c_0 N^{-\beta}$.

Let $N_0 \geq 1$ and choose $\phi \equiv \frac{3}{4}\delta$ on D_{N_0}. Then $m_\Delta(\frac{1}{N}\phi_N \geq \frac{3}{4}\delta) \geq \mu_\Delta(E_N) \geq c_0 N^{-\beta}$ for all $N \geq N_0$. Set $\phi \equiv -a$ on $\Delta - D_{N_0}$ where $a > 0$ is chosen so that $\bar{\phi} = 0$. Since $a \to 0$ as $N_0 \to \infty$, we can choose N_0 sufficiently large that $\|\phi\| = \frac{3}{4}\delta$. Hence the required estimates hold for ϕ and for all nearby observables. ∎

Let $c_1 = c_0/2$. Let $\mathcal{A}$ denote the subset of Hölder observables ψ satisfying the property that there exists $\epsilon_0 > 0$ such that for all $\epsilon \in (0, \epsilon_0)$ we have $\mu(|\frac{1}{N}\psi_N - \bar{\psi}| \geq \epsilon) \geq c_1/N^\beta$ for infinitely many values of N.

Corollary A.2 *The subset $\mathcal{A}$ is open and dense in the space of Hölder observables.*

Proof Suppose that $\psi \notin \mathcal{A}$. Given $\delta > 0$ construct ϕ as above and let $\psi' = \psi + \phi$. (So $\bar{\psi}' = \bar{\psi}$.) Choose $\epsilon_0 < \delta/4$. Then there exists $\epsilon \in (0, \epsilon_0)$ such that $\mu(|\frac{1}{N}\psi_N - \bar{\psi}| \geq \epsilon) \leq c_1/N^\beta$ for all sufficiently large N. Since $\delta > 4\epsilon$, $\mu(|\frac{1}{N}\phi_N| \geq 2\epsilon) \geq c_0/N^\beta$ for sufficiently large N. It follows that

$$\begin{aligned}\mu(|\tfrac{1}{N}\psi'_N - \bar{\psi}'| \geq \epsilon) &\geq \mu(|\tfrac{1}{N}\phi_N| \geq 2\epsilon \text{ and } |\tfrac{1}{N}\psi_N - \bar{\psi}| < \epsilon) \\ &\geq \mu(|\tfrac{1}{N}\phi_N| \geq 2\epsilon) - \mu(|\tfrac{1}{N}\psi_N - \bar{\psi}| \geq \epsilon) \geq (c_0 - c_1)/N^\beta = c_1/N^\beta.\end{aligned}$$

Since ϵ_0 (and hence ϵ) is arbitrarily small, the result follows. ∎

Acknowledgements This research was supported in part by EPSRC Grant EP/D055520/1. I am very grateful to Matthew Nicol for helpful discussions and to the University of Houston for the use of e-mail.

References

[1] V. Araújo. Large deviations bound for semiflows over a non-uniformly expanding base. *Bull. Braz. Math. Soc. (N.S.)* **38** (2007) 335–376.

[2] V. Araújo and M. J. Pacifico. Large deviations for non-uniformly expanding maps. *J. Stat. Phys.* **125** (2006) 415–457.

[3] N. Chernov and R. Markarian. Dispersing billiards with cusps: slow decay of correlations. *Comm. Math. Phys.* **270** (2007) 727–758.

[4] N. I. Chernov and H.-K. Zhang. Billiards with polynomial mixing rates. *Nonlinearity* **18** (2005) 1527–1553.

[5] N. I. Chernov and H.-K. Zhang. Improved estimates for correlations in billiards. *Commun. Math. Phys.* **77** (2008) 305–321.

[6] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, second ed., Applications of Mathematics **38**, Springer-Verlag, New York, 1998.

[7] R. S. Ellis. *Entropy, large deviations, and statistical mechanics.* Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] **271**, Springer-Verlag, New York, 1985.

[8] H. Hennion and L. Hervé. *Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness.* Lecture Notes in Math. **1766**, Springer, Berlin, 2001.

[9] H. Hu. Decay of correlations for piecewise smooth maps with indifferent fixed points. *Ergodic Theory Dynam. Systems* **24** (2004) 495–524.

[10] G. Keller and T. Nowicki. Spectral theory, zeta functions and the distribution of periodic points for Collet-Eckmann maps. *Comm. Math. Phys.* **149** (1992) 31–69.

[11] Y. Kifer. Large deviations in dynamical systems and stochastic processes. *Trans. Amer. Math. Soc.* **321** (1990) 505–524.

[12] C. Liverani, B. Saussol, and S. Vaienti. A probabilistic approach to intermittency. *Ergodic Theory Dynam. Systems* **19** (1999) 671–685.

[13] A. Lopes. Entropy and large deviation. *Nonlinearity* **3** (1990) 527–546.

[14] R. Markarian. Billiards with polynomial decay of correlations. *Ergodic Theory Dynam. Systems* **24** (2004) 177–197.

[15] I. Melbourne and M. Nicol. Large deviations for nonuniformly hyperbolic systems. *Trans. Amer. Math. Soc.* **360** (2008) 6661–6676.

[16] F. Merlevède, M. Peligrad, and S. Utev. Recent advances in invariance principles for stationary sequences. *Probab. Surv.* **3** (2006) 1–36 (electronic).

[17] S. Orey and S. Pelikan. Large deviation principles for stationary processes. *Ann. Probab.* **16** (1988) 1481–1495.

[18] M. Pollicott and R. Sharp. Large deviations for intermittent maps. Preprint, 2008.

[19] M. Pollicott, R. Sharp and M. Yuri. Large deviations for maps with indifferent fixed points, *Nonlinearity* **11** (1998) 1173–1184.

[20] Y. Pomeau and P. Manneville. Intermittent transition to turbulence in dissipative dynamical systems. *Comm. Math. Phys.* **74** (1980) 189–197.

[21] L. Rey-Bellet and L.-S. Young. Large deviations in non-uniformly hyperbolic dynamical systems. *Ergodic Theory Dynam. Systems* **28** (2008) 587–612.

[22] E. Rio. *Théorie asymptotique des processus aléatoires faiblement dépendants.* Mathématiques & Applications **31**, Springer-Verlag, Berlin, 2000.

[23] S. Waddington. Large deviation asymptotics for Anosov flows. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **13** (1996) 445–484.

[24] L.-S. Young. Large deviations in dynamical systems. *Trans. Amer. Math. Soc.* **318** (1990) 525–543.

[25] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.* **147** (1998) 585–650.

[26] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.* **110** (1999) 153–188.