Decay of correlations for nonuniformly expanding systems with general return times

Ian Melbourne [*] Dalia Terhesiu [†]

20 May, 2011. Updated 2 October, 2012

Abstract

We give a unified treatment of decay of correlations for nonuniformly expanding systems with a good inducing scheme. In addition to being more elementary than previous treatments, our results hold for general integrable return time functions under fairly mild conditions on the inducing scheme.

1 Introduction

Let $T : X \to X$ be a (noninvertible) measure preserving transformation with ergodic invariant probability measure μ_X. Given $v \in L^1(X)$, $w \in L^\infty(X)$, we define the correlation function $\rho_{v,w}(n) = \int_X v\, w \circ T^n \, d\mu_X - \int_X v \, d\mu_X \int_X w \, d\mu_X$. If T is mixing, then $\rho_{v,w}(n) \to 0$ as $n \to \infty$.

Definition 1.1 Let $\mathcal{B}(X) \subset L^1(X)$ denote a collection of observables $v : X \to \mathbb{R}$. Let $a_n > 0$ be a real sequence with $a_n \to 0$. We say that T has *uniform decay rate* a_n for observables in $\mathcal{B}(X)$ if for every $v \in \mathcal{B}(X)$ there is a constant $C_v > 0$ such that $|\rho_{v,w}(n)| \le C_v |w|_\infty a_n$ for all $w \in L^\infty(X)$.

We assume the existence of an induced map $F : Y \to Y$, $Y \subset X$, given by $F(y) = f^{\varphi(y)}(y)$ for some return time $\varphi : Y \to \mathbb{Z}^+$. (We do not require that φ is the first return time to Y.) It is assumed throughout that μ is an F-invariant ergodic probability measure on Y and that $\varphi \in L^1(Y)$. The measure μ_X on X is constructed from μ and φ in the standard way (see Section 2.1). The idea is to recover decay properties for T from properties of F and the return tails $\mu(y \in Y : \varphi(y) > n)$.

[*]Department of Mathematics, University of Surrey, Guildford, Surrey GU2 7XH, UK. Current address: Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK

[†]Department of Mathematics, University of Surrey, Guildford, Surrey GU2 7XH, UK. Current address: University of Rome Tor Vergata, via della Ricerca Scientifica, I-00133, Rome, Italy

In this paper, we combine the method of operator renewal sequences [5, 6, 17] with dynamical truncation [13] to give a particularly elementary and general treatment of decay of correlations in a much wider context than the usual Young tower setting [19]. Moreover, our results are strictly sharper than those obtained in the setting of Young towers by the methods of coupling [19], Birkhoff cones [12] and stochastic perturbation [10].

1.1 Young towers

Young [18, 19] considered the case where $T : X \to X$ is an ergodic nonuniformly expanding local diffeomorphism on a manifold X modelled by a *Young tower*. In particular, $F : Y \to Y$ is a uniformly expanding map with good distortion properties with respect to a countable partition (a so-called Gibbs-Markov map) and φ is constant on partition elements. Throughout this subsection, we take $\mathcal{B}(X)$ to be the space of (piecewise) Hölder observables.

In the case where $\mu(\varphi > n)$ decays exponentially, Young [19] obtained exponential decay of correlations. In the subexponential case, Young [19] proved (amongst other things) that if $\mu(\varphi > n) = O(1/n^{\beta+1})$, $\beta > 0$, then correlations decay at the rate $a_n = 1/n^{\beta}$. This result was shown to be optimal by Sarig [17] and Gouëzel [5]. Gouëzel [6] introduced a very general class of *convolutive sequences* and proved that if $\mu(\varphi > n) = O(b_n)$ where b_n is convolutive, then decay of correlations holds with optimal rate $a_n = \sum_{j>n} b_j$. This includes the cases of stretched exponential decay of correlations (Example 5.6) and *polynomially decreasing* sequences (Example 5.1).

Even in the context of Young towers, we obtain a number of new results. We mention three of these now. (The general formulation Theorem 4.2 of our result is somewhat technical and hence delayed until Section 4.)

Theorem 1.2 *Suppose that $\varphi \in L^{1+\epsilon}(Y)$ for some $\epsilon > 0$. Then for any $p > 0$ there exists $\delta > 0$, $C > 0$ such that*

$$\rho_{v,w}(n) \le C\|v\||w|_\infty \Big\{ \sum_{j>\delta n} \mu(\varphi > j) + n\mu(\varphi > \delta n) + O(n^{-p}) \Big\},$$

for all $v \in \mathcal{B}(X)$, $w \in L^\infty(X)$, $n \ge 1$.

An immediate consequence of Theorem 1.2 is optimal upper bounds on decay of correlations when $\mu(\varphi > n) = O(1/n^{\beta+1})$ for $\beta > 0$. More generally, the case when $\mu(\varphi > n)$ is dominated by a regularly varying sequence $\ell(n)/n^{\beta+1}$ also follows from Theorem 1.2, as does the even more general situation where $\mu(\varphi > n)$ is dominated by a polynomially decreasing sequence. These results are stated in Section 5 along with treatments of exponential decay, stretched exponential decay, and regularly varying sequences with $\beta = 0$.

Next, we mention two theoretical results. It has been noted elsewhere that either (i) $\varphi \in L^2(Y)$, equivalently $\sum_{n=1}^\infty n\mu(\varphi > n) < \infty$, or (ii) summable decay of correlations $\sum_{n=1}^\infty \rho(n) < \infty$, are sufficient to guarantee the validity of the central limit theorem. The special case $q = 1$ of Corollary 1.3 below states that $\varphi \in L^2(Y)$ implies summable decay of correlations.

Corollary 1.3 *Let $q > 0$. If $\varphi \in L^{q+1}(Y)$, then $\sum_{n=1}^\infty n^{q-1}\rho_{v,w}(n) < \infty$ for all $v \in \mathcal{B}(X)$, $w \in L^\infty(X)$.*

Proof By Theorem 1.2, we can choose $\delta > 0$ so that

$$\rho_{v,w}(n) \ll \sum_{j>\delta n} \mu(\varphi > j) + n\mu(\varphi > \delta n) + n^{-(q+1)}.$$

(Throughout, we use 'big O' and $\ll$ notation interchangeably, writing $a_N = O(b_N)$ or $a_N \ll b_N$ (as $N \to \infty$) if there is a constant $C > 0$ such that $a_N \le Cb_N$ for all $N \ge 1$.)

Multiplying by n^{q-1}, the last term is summable and the middle term yields $\sum_{n=1}^\infty n^q\mu(\varphi > \delta n) \ll \sum_{n=1}^\infty n^q\mu(\varphi > n) < \infty$. Finally,

$$\begin{aligned}\sum_{n=1}^\infty n^{q-1}\sum_{j>\delta n}\mu(\varphi > j) \ll \sum_{n=1}^\infty n^{q-1}\sum_{j>n}\mu(\varphi > j) = \sum_{j=2}^\infty\sum_{n<j} n^{q-1}\mu(\varphi > j) \\ \le \sum_{j=2}^\infty j^q\mu(\varphi > j) < \infty,\end{aligned}$$

so that $\sum_{n=1}^\infty n^{q-1}\rho_{v,w}(n) < \infty$. ∎

Our main results, including Theorem 1.2, give conditions for uniform rates of decay. A natural question is to inquire when uniform decay rates exist in the first place. The following result addresses this issue.

Theorem 1.4 *Suppose that $\mu(\varphi > n) = O((n\log n)^{-1})$. (We continue to assume in addition that $\varphi \in L^1(Y)$.) Then correlations decay at a uniform rate for $v \in \mathcal{B}(X)$, $w \in L^\infty(X)$.*

1.2 Systems with excellent inducing schemes

Let $T : X \to X$ be a transformation with induced map $F = f^\varphi : Y \to Y$ and F-invariant ergodic probability measure μ. Let $R : L^1(Y) \to L^1(Y)$ be the transfer operator for F, so $\int_Y Rv\,w\,d\mu = \int_Y v\,w\circ F\,d\mu$ for all $v \in L^1(Y)$, $w \in L^\infty(Y)$. Define $R_n = R1_{\{\varphi=n\}}$, $n \ge 1$.

Let $\mathcal{B}(Y) \subset L^1(Y)$ be a Banach space with norm $\|\;\|$ satisfying $|v|_1 \le \|v\|$ for all $v \in \mathcal{B}(Y)$, and such that constant functions lie in $\mathcal{B}(Y)$.

(H1) The operator $R_n : \mathcal{B}(Y) \to \mathcal{B}(Y)$ is bounded for all n, and $\|R_n\| \ll \mu(\varphi = n)$.

Let $\mathbb{D}$ and $\bar{\mathbb{D}}$ denote the open and closed unit disk in $\mathbb{C}$. Define $R(z) = \sum_{n=1}^{\infty} R_n z^n$ for $z \in \bar{\mathbb{D}}$. Hypothesis (H1) guarantees that $z \mapsto R(z)$ is a continuous family of bounded operators on $\mathcal{B}(Y)$ for $z \in \bar{\mathbb{D}}$, and the family is analytic on $\mathbb{D}$. Note that $R(1) = R$, so in particular 1 lies in the spectrum of $R(1)$.

(H2) (i) The eigenvalue 1 is simple and isolated in the spectrum of $R(1)$.

(ii) For $z \in \bar{\mathbb{D}} \setminus \{1\}$, the spectrum of $R(z)$ does not contain 1.

Definition 1.5 An inducing scheme $F = T^\varphi : Y \to Y$ is *excellent* if hypotheses (H1) and (H2) are satisfied for an appropriate Banach space $\mathcal{B}(Y)$.

Let $v : X \to \mathbb{R}$ be an observable. We say that $\hat{v} : Y \to \mathbb{R}$ is *derived* from v if for every $n \geq 1$, there exists $j \in \{1, \ldots, n-1\}$ such that $\hat{v}(y) = v(T^j y)$ for all $y \in Y$ with $\varphi(y) = n$. Let $\mathcal{D}_v$ denote the set of observables $\hat{v} : Y \to \mathbb{R}$ derived from v.

Definition 1.6 An observable $v : X \to \mathbb{R}$ is *exchangeable* if $\mathcal{D}_v$ is a bounded subset of $\mathcal{B}(Y)$. Set $\|v\| = \sup_{\hat{v} \in \mathcal{D}_v} \|\hat{v}\|$.

The definition of exchangeability formalises the need for control of iterates $T^j y$ for $j \in \{1, \ldots, \varphi(y) - 1\}$.

For excellent inducing schemes and exchangeable observables, we obtain almost identical results as those for Hölder observables on systems modelled by Young towers. In particular, Theorems 1.2 and 1.4 and Corollary 1.3 hold in this generality. If we assume further that $\mathcal{B}(Y)$ is embedded in $L^\infty(Y)$ (rather than in $L^1(Y)$), then all of our conclusions in this paper are identical to those for Young towers.

Example 1.7 (Young towers) The inducing schemes for the nonuniformly expanding maps studied by Young [18, 19] are Gibbs-Markov. These are excellent inducing schemes since it is well-known that hypotheses (H1) and (H2) are satisfied for the Banach space $\mathcal{B}(Y)$ consisting of piecewise Hölder observables on Y. Moreover, piecewise Hölder observables on X are exchangeable.

Although most of this paper is concerned with nonuniformly expanding maps, the results extend to systems that are nonuniformly hyperbolic in the sense of Young [18, 19]. Details of this extension are given in Appendix B based on ideas of [3, 7].

Example 1.8 (AFN maps) Zweimüller [20] studied a class of non-Markovian uniformly expanding interval maps (so-called AFN maps) with finite absolutely continuous invariant measures. In particular, [20] obtained a spectral decomposition into basic ergodic sets and proved that for each basic set there is a unique absolutely continuous invariant probability measure. Each basic set is mixing up to a finite cycle, and we suppose that X is a mixing basic set. There is a first return map $F : Y \to Y$

that is uniformly expanding with respect to a partition consisting of intervals. Moreover F has good distortion properties. It can be shown that F is an excellent inducing scheme with function space $\mathrm{BV}(Y)$ (observables of bounded variation).

Unfortunately, $\mathrm{BV}(X)$ is not exchangeable. However, it turns out that F is also excellent if we enlarge $\mathcal{B}(Y)$ to consist of piecewise bounded variable observables, and then the corresponding space $\mathcal{B}(X)$ is exchangeable. The details are sketched in Section 5.3.

Remark 1.9 Suppose that the inducing scheme is a first return map (that is, $\varphi(y) = \inf\{n \geq 1 : T^n y \in Y\}$). If $v : X \to \mathbb{R}$ is supported on Y and $1_Y v \in \mathcal{B}(Y)$, then v is exchangeable. Hence our results apply to such observables (and all $w \in L^\infty(X)$) whenever the first return map is an excellent inducing scheme.

1.3 Systems with good inducing schemes

There are a number of situations where the induced map has good behaviour but properties such as bounded distortion and/or large images fail. Examples include the class of interval maps studied by Araújo *et al.* [2] (where the induced map is of the type studied by Rychlik [16]), and Hu-Vaienti maps [9] which are multidimensional nonMarkovian nonuniformly expanding maps with indifferent fixed points.

In such situations, it is likely that hypothesis (H1) can fail quite badly. However, it turns out that we can obtain decay estimates (often optimal estimates) under a weaker condition (hypothesis (†) below) that seems much more tractable. See [11] for an application of our methods to the interval maps studied by [2].

Fix the Banach space $\mathcal{B}(Y)$ as before. We replace hypothesis (H1) by:

(†) $\sum_{n=1}^\infty \sum_{j>n} \|R_j\| < \infty$.

Note that $\mu(\varphi = n) = |R_n 1_Y|_1 \leq \|R_n\| \|1_Y\|$ so for excellent inducing schemes condition (†) is simply the requirement that $\varphi \in L^1(Y)$. In general, condition (†) is sufficient to ensure that the family $R(z) = \sum_{n=1}^\infty R_n z^n$ is analytic on $\mathbb{D}$ and continuous on $\bar{\mathbb{D}}$ as before. (The full strength of (†) is required in Proposition 3.4.)

Definition 1.10 An inducing scheme $F = T^\varphi : Y \to Y$ is *good* if hypotheses (†) and (H2) are satisfied for an appropriate Banach space $\mathcal{B}(Y)$.

We have the following generalisations of Theorem 1.2 and Corollary 1.3.

Theorem 1.11 *Suppose that $F = T^\varphi : Y \to Y$ is a good inducing scheme and that $\mathcal{B}(X)$ is a collection of exchangeable observables. Suppose further that $\sum_{n=1}^\infty n^\epsilon \sum_{j>n} \|R_j\| < \infty$ for some $\epsilon > 0$.*

Then for any $p > 0$ there exists $\delta > 0$, $C > 0$ such that

$$\rho_{v,w}(n) \leq C\|v\||w|_\infty \Big\{ \sum_{j>\delta n} \mu(\varphi > j) + n\mu(\varphi > \delta n) + O(n^{-p}) \Big\},$$

for all $v \in \mathcal{B}(X)$, $w \in L^\infty(X)$, $n \ge 1$.

Corollary 1.12 *Suppose that $F = T^\varphi : Y \to Y$ is a good inducing scheme and that $\mathcal{B}(X)$ is a collection of exchangeable observables. Suppose further that $\sum_{n=1}^\infty n^\epsilon \sum_{j>n} \|R_j\| < \infty$ for some $\epsilon > 0$. Let $q > 0$. If $\varphi \in L^{q+1}(Y)$, then $\sum_{n=1}^\infty n^{q-1}\rho_{v,w}(n) < \infty$ for all $v \in \mathcal{B}(X)$, $w \in L^\infty(X)$.* ∎

The remainder of this paper is as follows. In Section 2, we describe the strategy adopted in this paper. In essence, everything that follows Section 2 is an extended exercise. The required estimates are carried out in Sections 3 and 4. In particular, Section 4 contains the most general versions of our results. In Section 5, we verify that Theorem 1.11 and Theorem 1.4 follow from the general results and compute correlation decay rates for specific tail functions $\mu(\varphi > n)$.

Remark 1.13 The technique introduced in this paper can also be used to obtain a simplified and generalised treatment of lower bounds (and improved upper bounds) for decay of correlations [5, 6, 17]. The results on lower bounds are restricted to the setting of excellent first return maps and observables supported on Y. Since the setting is more restricted, and additional ideas are required, we defer these results to a later paper.

2 Strategy

The strategy in this paper consists of three main steps:

1. Pass to a *tower extension* $f : \Delta \to \Delta$ of the underlying map $T : X \to X$. The tower Δ is a discrete suspension over $F : Y \to Y$ with height φ. In particular $F = T^\varphi = f^\varphi$. Decay of correlations on Δ pushes down to decay of correlations on X. Hence this step reduces to the situation where φ is a *first* return time function.

2. Use *dynamical truncation* [13] to replace the tower Δ by a tower Δ^* with finite height φ^* in such a way that the first return map F is unchanged. The truncation error between correlation decay on Δ and on Δ^* is easily controlled.

3. Use *operator renewal sequences* [5, 6, 17] to estimate correlation decay on the truncated tower Δ^* in terms of the height φ^* and spectral properties of the transfer operator R for the induced map $F : Y \to Y$. A key observation from [13] is that the dependence of the estimates on φ^* are explicit, while $F : Y \to Y$, R and μ are unchanged throughout.

We now describe each of these steps in more detail.

2.1 Tower extension

Given the induced map $F : Y \to Y$ and return time $\varphi : Y \to \mathbb{Z}^+$, we define the tower $\Delta = Y^\varphi = \{(y, \ell) \in Y \times \mathbb{Z} : 0 \le \ell \le \varphi(y) - 1\}$ and the tower map $f : \Delta \to \Delta$ by $f(y, \ell) = (y, \ell + 1)$ for $\ell \le \varphi(y) - 2$ and $f(y, \varphi(y) - 1) = (Fy, 0)$. Set $\bar\varphi = \int_Y \varphi \, d\mu$ and define the f-invariant probability measure $\mu_\Delta = (\mu \times \text{counting})/\bar\varphi$ on Δ.

Define the semiconjugacy $\pi : \Delta \to X$, $\pi(y, \ell) = T^\ell y$, and set $\mu_X = \pi_*\mu_\Delta$. Given observables $v, w : X \to \mathbb{R}$, we define the lifted observables $v \circ \pi, w \circ \pi : \Delta \to \mathbb{R}$. Then it suffices to compute correlation decay rates for the lifted observables on Δ. Moreover, it is immediate that if $v : X \to \mathbb{R}$ is exchangeable (relative to the Banach space $\mathcal{B}(Y)$) then so is the lifted observable $v \circ \pi : \Delta \to \mathbb{R}$.

From now on, given $v, w : \Delta \to \mathbb{R}$, we study decay rates for

$$\rho_{v,w}(n) = \int_\Delta v \, w \circ f^n \, d\mu_\Delta - \int_\Delta v \, d\mu_\Delta \int_\Delta w \, d\mu_\Delta.$$

2.2 Dynamical truncation

Given $k \ge 1$, we define the truncated return time function $\varphi^* = \min\{\varphi, k\}$. Just as we defined $f : \Delta \to \Delta$ starting from $F : Y \to Y$ and $\varphi : Y \to \mathbb{Z}^+$, we can define the truncated tower map $f^* : \Delta^* \to \Delta^*$ starting from $F : Y \to Y$ and $\varphi^* : Y \to \mathbb{Z}^+$. Note that $F = f^\varphi = (f^*)^{\varphi^*}$ is independent of k.

Similarly, set $\bar\varphi^* = \int_Y \varphi^* \, d\mu$ and define the f^*-invariant probability measure $\mu_{\Delta^*} = (\mu \times \text{counting})/\bar\varphi^*$ on Δ^*.

Given $v \in L^\infty(\Delta)$, $w \in L^\infty(\Delta)$, we define $v \in L^\infty(\Delta^*)$, $w \in L^\infty(\Delta^*)$ by restriction. Let

$$\rho^*_{v,w}(n) = \int_{\Delta^*} v \, w \circ (f^*)^n \, d\mu_{\Delta^*} - \int_{\Delta^*} v \, d\mu_{\Delta^*} \int_{\Delta^*} w \, d\mu_{\Delta^*}.$$

We have the estimate [13],

$$|\rho_{v,w}(n) - \rho^*_{v,w}(n)| \ll |v|_\infty |w|_\infty \Big(\sum_{j > k} \mu(\varphi > j) + n\mu(\varphi > k) \Big). \tag{2.1}$$

See Appendix A for details.

2.3 Operator renewal sequences

It remains to estimate decay of correlations on the truncated tower. Since φ^* is bounded, we expect to obtain an exponential estimate of the form $|\rho^*_{v,w}(n)| \le C_{v,w}(k) e^{-a(k)n}$. Given sufficient control of $C(k)$ and $a(k)$, this estimate can be combined with (2.1) (choosing $k = k(n)$) to obtain an estimate for $\rho_{v,w}(n)$. A surprising aspect of our approach is the degree of control on $C(k)$ and $a(k)$.

We recall the standard definitions of renewal theory, first for the nontruncated map. Let L denote the transfer operator for $f : \Delta \to \Delta$ and let R denote the transfer operator for $F = f^\varphi : Y \to Y$. Define the renewal operators $T_n, R_n : \mathcal{B}(Y) \to \mathcal{B}(Y)$

$$T_n = 1_Y L^n 1_Y,\ \ n \ge 0, \quad R_n = 1_Y L^n 1_{\{\varphi=n\}} = R1_{\{\varphi=n\}},\ \ n \ge 1.$$

Define $T(z) = \sum_{n=0}^\infty T_n z^n$ and $R(z) = \sum_{n=1}^\infty R_n z^n$. An elementary calculation shows that $T_n = \sum_{j=1}^n T_{n-j} R_j$ and hence $T(z) = I + T(z)R(z)$ leading to the renewal equation $T(z) = (I - R(z))^{-1}$. Hypothesis (H1) or (†) guarantees that $R(z)$ is analytic on $\mathbb{D}$ and continuous on $\bar{\mathbb{D}}$. Moreover, $T(z)$ is analytic on $\mathbb{D}$ and It follows from (H2)(ii) that $T(z)$ extends continuously to $\bar{\mathbb{D}} \setminus \{1\}$. By (H2)(i), $T(z)$ has a singularity at $z = 1$. The idea of renewal sequences is to use knowledge about the sequence R_n and the singularity to understand the behaviour of $T(z)$ and thereby T_n (and ultimately L^n).

The situation is simpler for the truncated dynamical system. Passing to the truncated tower, we have the transfer operator L^* corresponding to $f^* : \Delta^* \to \Delta^*$. By construction the first return map $F = (f^*)^{\varphi^*} : Y \to Y$ is independent of k with fixed transfer operator R. Define the truncated renewal operators

$$T_n^* = 1_Y L^{*n} 1_Y,\ \ n \ge 0, \quad R_n^* = 1_Y L^{*n} 1_{\{\varphi^*=n\}} = R1_{\{\varphi^*=n\}},\ \ n \ge 1.$$

Again, $T^*(z) = \sum_{n=0}^\infty T_n^* z^n$ is analytic on $\mathbb{D}$. Evidently, $R_n^* = 0$ for $n > k$, so $R^*(z) = \sum_{n=1}^k R_n^* z^n$ is a polynomial. Again, we have the renewal equation $T^*(z) = (I - R^*(z))^{-1}$. For the truncated tower, it follows from standard arguments that the singularity of $T^*(z)$ at $z = 1$ is a simple pole.

In Section 3, we investigate the behaviour of $T^*(z)$ using the ideas described above. In Section 4, we show how to pass from $T^*(z)$ to $L^*(z) = \sum_{n=0}^\infty L^{*n} z^n$. From this we obtain exponential convergence results for the coefficients L^{*n} and hence the required exponential decay for $\rho^*_{v,w}(n)$.

3 Analyticity of $T^*(z)$

In this section, we assume that we have a good inducing scheme $F : Y \to Y$ with transfer operator R satisfying conditions (†) and (H2) for an appropriate Banach space $\mathcal{B}(Y) \subset L^1(Y)$. Denote the spectral projection corresponding to the simple eigenvalue 1 for $R(1)$ by $Pv = \int_Y v\, d\mu$.

For $a \ge 0$, let $\mathbb{D}_a = \{|z| \in \mathbb{C} : |z| < e^a\}$. Define

$$S_q(k,a) = \sum_{j=1}^k \Big(\sum_{\ell > j} \|R_\ell\|\Big) j^q e^{ja}, \quad k \ge 1,\ \ a, q \in [0, \infty).$$

We prove the following result.

Lemma 3.1 *Let $a = a(k) \in (0,\infty)$ be such that $\lim_{k\to\infty} a^r S_r(k,a) = 0$ for some $r \in (0,1]$. Then there exists $k_0 \geq 1$ such that for any $q \in (0,1]$, $k \geq k_0$,*

$$T^*(z) = (1-z)^{-1}(1/\bar{\varphi}^*)P + J^*(z),$$

where $J^*(z)$ is analytic on the disk $\mathbb{D}_a$ and $\sup_{z\in\mathbb{D}_a} |z-1|^{1-q}\|J^*(z)\| \ll S_q(k,a)$.

In the remainder of this section, we prove Lemma 3.1. As already mentioned, it follows from standard arguments that $T^*(z)$ has a simple pole at $z = 1$ and so $B^*(z) = (1-z)T^*(z)$ extends analytically to $\mathbb{D}_a$ for some $a > 0$. The proof of Lemma 3.1 consists of estimating $a = a(k)$ and controlling the norms of various analytic families of operators on $\mathbb{D}_a$. This is a fairly routine exercise, but the calculations are quite complicated. To remedy this, we first sketch the formal calculation in Subsection 3.1 and then carry out the rigorous estimates in Subsection 3.2.

3.1 Formal calculation on $\bar{\mathbb{D}}$

In this subsection, we regard k as fixed and large, and we argue formally. Note that $R^*(1) = R(1)$ with simple isolated eigenvalue 1. Moreover $R^*(z)$ is a polynomial, so there exists $\delta > 0$ such that the eigenvalue 1 for $R^*(1)$ extends to an analytic family of eigenvalues $\lambda^*(z)$ on $B_\delta(1)$ with a corresponding family of spectral projections $P^*(z)$. Let $Q^*(z) = I - P^*(z)$. Then in an obvious notation, we can write

$$T^*(z) = (1-\lambda^*(z))^{-1}P^*(z) + (I - R^*(z))^{-1}Q^*(z),$$

for $z \in \bar{\mathbb{D}} \cap B_\delta(1)$, $z \neq 1$. A standard calculation (eg. [15]) shows that $\lambda^*(z) = 1 + (z-1)\bar{\varphi}^* + O(|z-1|^2)$ and hence $T^*(z)$ has a pole of order 1 at $z = 1$. In particular, the function $B^*(z) = (1-z)T^*(z)$ is analytic on $\mathbb{D}_a$ for some $a > 0$. Moreover, $B^*(1) = (1/\bar{\varphi}^*)P$. Thus we can write $B^*(z) = (1/\bar{\varphi}^*)P + (1-z)J^*(z)$ where $J^*(z)$ is analytic on $\mathbb{D}_a$. Now divide by $(1-z)$ to obtain the formula for $T^*(z)$ in Lemma 3.1.

We end this subsection by deriving a formula for $J^*(z)$. Write

$$P^*(z) = P + (z-1)P_1^*(z), \qquad \lambda^*(z) = 1 + (z-1)\{\bar{\varphi}^* + \tilde{\lambda}^*(z)\},$$

where $\tilde{\lambda}^*(1) = 0$. Then (at least formally),

$$\Big(\frac{1-\lambda^*(z)}{1-z}\Big)^{-1} = \frac{1}{\bar{\varphi}^*}\Big(1 + \frac{1}{\bar{\varphi}^*}\tilde{\lambda}^*(z)\Big)^{-1} = \frac{1}{\bar{\varphi}^*} - \Big(\frac{1}{\bar{\varphi}^*}\Big)^2\tilde{\lambda}^*(z)\Big(1 + \frac{1}{\bar{\varphi}^*}\tilde{\lambda}^*(z)\Big)^{-1},$$

and hence

$$J^*(z) = \begin{cases} (z-1)^{-1}(1/\bar{\varphi}^*)^2\tilde{\lambda}^*(z)\big\{1 + (1/\bar{\varphi}^*)\tilde{\lambda}^*(z)\big\}^{-1}P & \\ \qquad -(1/\bar{\varphi}^*)\big\{1 + (1/\bar{\varphi}^*)\tilde{\lambda}^*(z)\big\}^{-1}P_1^*(z) & \\ \qquad +(I - R^*(z))^{-1}Q^*(z), & z \in \mathbb{D}_a \cap B_\delta(1) \\ (I - R^*(z))^{-1} - (1-z)^{-1}(1/\bar{\varphi}^*)P, & z \in \mathbb{D}_a \setminus B_\delta(1) \end{cases} \tag{3.1}$$

3.2 Rigorous calculation on $\mathbb{D}_a$

By (H2)(i), we can choose a closed loop $\Gamma \in \mathbb{C} \setminus \operatorname{spec} R(1)$ separating 1 from the remainder of the spectrum of $R(1)$. There exists $\delta > 0$ such that the spectrum of $R(z)$ does not intersect Γ for $z \in \bar{\mathbb{D}} \cap B_\delta(1)$ and we can define the spectral projection

$$P(z) = \frac{1}{2\pi i} \int_\Gamma (\xi I - R(z))^{-1} \, d\xi. \tag{3.2}$$

For $z \in \bar{\mathbb{D}} \cap B_\delta(1)$, define the corresponding eigenvalue $\lambda(z)$, so $R(z)P(z) = \lambda(z)P(z)$, and the complementary projection $Q(z) = I - P(z)$.

For k sufficiently large, and z close enough to 1, we can define similarly $\lambda^*(z)$, $P^*(z)$ and $Q^*(z)$. The next result is a uniform version of this statement.

Proposition 3.2 *Suppose that $a = a(k)$ satisfies $\lim_{k\to\infty} aS_0(k,a) = 0$. Then*

(a) For any $\delta > 0$, there exists $k_0 \geq 1$ such that $\|(I - R^*(z))^{-1}\| \ll 1$ for $k \geq k_0$, $z \in \mathbb{D}_a \setminus B_\delta(1)$.

(b) There exists $\delta > 0$ and $k_0 \geq 1$ such that for all $k \geq k_0$ there exists a continuous family $z \mapsto \lambda^*(z)$, $z \in \mathbb{D}_a \cap B_\delta(1)$, of simple eigenvalues for $R^*(z)$ satisfying $\lambda^*(1) = 1$. Moreover, $|\lambda^*(z)| \ll 1$ for $k \geq k_0$, $z \in \mathbb{D}_a \cap B_\delta(1)$.

(c) The spectral projections $P^*(z)$ corresponding to the eigenvalues $\lambda^*(z)$ satisfy $\|P^*(z)\| \ll 1$ for $k \geq k_0$, $z \in \mathbb{D}_a \cap B_\delta(1)$.

(d) $\|(I - R^*(z))^{-1}Q^*(z)\| \ll 1$ for $k \geq k_0$, $z \in \mathbb{D}_a \cap B_\delta(1)$.

Proof We break the proof into three steps. First we work with $R(z)$, etc, on $\bar{\mathbb{D}}$. Second, we consider $R^*(z)$, etc, on $\bar{\mathbb{D}}$. Third, we consider $R^*(z)$, etc, on $\mathbb{D}_a$.
1.) By (†), $R(z)$ is uniformly convergent and hence continuous on $\bar{\mathbb{D}}$. Thus the results for $(I - R(z))^{-1}$, $\lambda(z)$, $P(z)$ and $(I - R(z))^{-1}Q(z)$ follow from (H2).
2.) Note that $\|R(z) - R^*(z)\| \leq 2\sum_{j>k} \|R_j\| \to 0$ as $k \to \infty$ uniformly on $\bar{\mathbb{D}}$ by (†). Hence the results for $(I - R^*(z))^{-1}$, $\lambda^*(z)$, $P^*(z)$ and $(I - R^*(z))^{-1}Q^*(z)$ on $\bar{\mathbb{D}}$ follow from step 1 and the resolvent identity.
3.) We claim that $\|R^*(e^{a+ib}) - R^*(e^a)\| \ll aS_0(k,a)$. By assumption, $aS_0(k,a) \to 0$ as $k \to 0$, so the result follows from step 2 and the resolvent identity.

To verify the claim, compute that

$$
\begin{aligned}
&\|R^*(e^{a+ib}) - R^*(e^{ib})\| \le \sum_{j=1}^{k} \|R_j^*\|(e^{ja}-1) \\
&= \sum_{j=1}^{k} (\sum_{\ell \ge j} \|R_\ell^*\|)(e^{ja}-1) - \sum_{j=1}^{k} (\sum_{\ell \ge j+1} \|R_\ell^*\|)(e^{ja}-1) \\
&= \sum_{j=1}^{k} (\sum_{\ell \ge j} \|R_\ell^*\|)(e^{ja}-1) - \sum_{j=1}^{k} (\sum_{\ell \ge j} \|R_\ell^*\|)(e^{(j-1)a}-1) \\
&= (e^a-1)\sum_{j=0}^{k-1} (\sum_{\ell > j} \|R_\ell^*\|)e^{ja} = (e^a-1)\sum_{j=0}^{k-1} (\sum_{\ell > j} \|R_\ell\|)e^{ja} = (e^a-1)S_0(k,a),
\end{aligned}
$$

as required. ∎

Define the polynomials of degree $k-1$,

$$
R_1^*(z) = \frac{R^*(z)-R^*(1)}{z-1}, \quad \tilde{R}^*(z) = \frac{R^*(z)-R^*(1)}{z-1} - \frac{dR^*}{dz}(1) = R_1^*(z) - R_1^*(1).
$$

Similarly, starting from P^* and λ^* instead of R^*, define the analytic functions $P_1^*(z)$, $\tilde{P}^*(z)$, $\lambda_1^*(z)$ and $\tilde{\lambda}^*(z)$.

Proposition 3.3 *Suppose that $a = a(k)$ satisfies $\lim_{k\to\infty} aS_0(k,a) = 0$. There exists $\delta > 0$ and $k_0 \ge 1$ such that for all $k \ge k_0$ and $z \in \mathbb{D}_a \cap B_\delta(1)$,*

- *(a)* $\|P_1^*(z)\| \ll \|R_1^*(z)\|$ *and* $|\lambda_1^*(z)| \ll \|R_1^*(z)\|$.
- *(b)* $|\tilde{\lambda}^*(z)| \ll \|\tilde{R}^*(z)\| + |z-1|\|R_1^*(z)\|^2$.

Proof (a) The estimate for P_1^* follows from equation (3.2) and the resolvent identity. Next,

$$
\begin{aligned}
(\lambda^*(z)-1)P &= \lambda^*(z)P^*(z) - \lambda^*(1)P^*(1) \; - \; \lambda^*(z)(P^*(z)-P^*(1)) \\
&= R^*(z)P^*(z) - R^*(1)P^*(1) \; - \; \lambda^*(z)(P^*(z)-P^*(1)),
\end{aligned}
$$

so the estimate for λ_1^* follows from the estimate for P_1^*.
(b) By equation (3.2), $\tilde{P}^*(z) = \frac{1}{2\pi i}\int_\Gamma (I + II)\, d\xi$, where

$$
\begin{aligned}
I &= (\xi I - R^*(1))^{-1}\tilde{R}^*(z)(\xi I - R^*(1))^{-1} \\
II &= \{(\xi I - R^*(z))^{-1} - (\xi I - R^*(1))^{-1}\}R_1^*(z)(\xi I - R^*(1))^{-1}.
\end{aligned}
$$

Then $\|I\| \ll \|\tilde{R}^*(z)\|$ and $\|II\| \ll |z-1|\|R_1^*(z)\|^2$. Hence $\|\tilde{P}^*(z)\| \ll \|\tilde{R}^*(z)\| + |z-1|\|R_1^*(z)\|^2$.

Next, define $S^*(z) = R^*(z)P^*(z)$ and correspondingly $\tilde{S}^*(z)$. Then

$$\tilde{S}^*(z) = R^*(z)\tilde{P}^*(z) + \tilde{R}^*(z)P^*(1) + (R^*(z) - R^*(1))\frac{dP^*}{dz}(1).$$

so that $\|\tilde{S}^*(z)\| \ll \|\tilde{R}^*(z)\| + |z-1|\|R_1^*(z)\|^2$. Finally,

$$\tilde{\lambda}^*(z)P = \tilde{S}^*(z) - \tilde{P}^*(z) - (z-1)^{-1}(\lambda^*(z) - \lambda^*(1))(P^*(z) - P^*(1)),$$

yielding the required estimate for $\tilde{\lambda}^*$. ∎

Proposition 3.4 *Let $r \in [0,1]$, $a > 0$, $b \in [0, 2\pi]$.*

(a) $\|R_1^*(e^{a+ib}) - R_1^*(e^{ib})\| \ll a^r S_r(k,a)$.

(b) $\|R_1^*(z)\| \ll 1 + a^r S_r(k,a)$ for all $z \in \mathbb{D}_a$.

(c) For any $\epsilon > 0$, there exists $\delta > 0$, $k_0 \geq 1$ such that $\|\tilde{R}^*(z)\| \leq \epsilon + a^r S_r(k,a)$ for all $k \geq k_0$ and $z \in \mathbb{D}_a \cap B_\delta(1)$.

Proof Write $U_j = \sum_{\ell > j} R_\ell$. The same calculation as the one used in the proof of Proposition 3.2 shows that $R_1^*(z) = \sum_{j=1}^{k} \sum_{\ell \geq j} R_\ell^* z^{j-1} = \sum_{j=0}^{k-1} U_j z^j$. Hence

$$\begin{aligned} \|R_1^*(e^{a+ib}) - R_1^*(e^{ib})\| &\leq \sum_{j=1}^{k-1} \|U_j\|(e^{ja} - 1) \leq a \sum_{j=1}^{M} j\|U_j\| e^{ja} + \sum_{j=M}^{k} \|U_j\| e^{ja} \\ &\leq aM^{1-r} \sum_{j=1}^{M} j^r \|U_j\| e^{ja} + M^{-r} \sum_{j=M}^{k} j^r \|U_j\| e^{ja} \end{aligned}$$

so taking $M \approx 1/a$ yields part (a).

By (†), R_1^* is bounded on $\bar{\mathbb{D}}$ uniformly in k. Hence (b) follows from (a).

Let $S(z) = \sum_{j=0}^{\infty} U_j z^j$ and note that S is absolutely summable on $\bar{\mathbb{D}}$ by (†). In particular, R_1^* and S are continuous on $\bar{\mathbb{D}}$. Moreover, S is independent of k and we can choose δ so that $\|S(z) - S(1)\| < \epsilon/2$ for $z \in \bar{\mathbb{D}} \cap B_\delta(1)$. Choose k_0 so that $\|S(z) - R_1^*(z)\| < \epsilon/4$ for all $z \in \bar{\mathbb{D}}$, $k \geq k_0$. Writing $\tilde{R}^*(z) = R_1^*(z) - R_1^*(1) = (S(z) - S(1)) - (S(z) - R_1^*(z)) + (S(1) - R_1^*(1))$ we obtain that $\|\tilde{R}^*\| < \epsilon$ for all $z \in \bar{\mathbb{D}} \cap B_\delta(1)$, $k \geq k_0$. Hence (c) follows from (a). ∎

Recall that the definition of J^* in (3.1) relied on the invertibility of $1 + (1/\bar{\varphi}^*)\tilde{\lambda}^*(z)$.

Corollary 3.5 *If $a^r S_r(k,a) \to 0$ for some $r \in (0,1]$, then there exists $k_0 \geq 1$ such that J^* is well-defined on $\mathbb{D}_a$ and $\|J^*(z)\| \ll |z-1|^{-1}\|\tilde{R}^*(z)\| + 1$ for all $k \geq k_0$ and $z \in \mathbb{D}_a$.*

Proof In particular, $aS_0(k,a) \to 0$, so Proposition 3.2 applies. Hence P^*, λ^*, and so on exist and are uniformly bounded on $\mathbb{D}_a$. By Propositions 3.3 and 3.4, $|\tilde\lambda^*| \ll \epsilon + a^r S_r(k,a) + (z-1)(1+aS_r(k,a))^2$, so choosing δ and k_0 appropriately, we can arrange that $(1/\bar\varphi^*)|\tilde\lambda^*| < \frac{1}{2}$ (say). Hence the formal expression for J^* makes sense. Moreover all terms in this expression are uniformly bounded except possibly for P_1^* and $(z-1)^{-1}\tilde\lambda^*$. By Propositions 3.3 and 3.4, $\|P_1^*\| \ll \|R_1^*\| \ll 1 + a^r S_r(k,a) \ll 1$, and $|(z-1)^{-1}\tilde\lambda^*| \ll |z-1|^{-1}\|\tilde R^*\| + \|R_1^*\|^2 \ll |z-1|^{-1}\|\tilde R^*\| + 1$. ∎

Proof of Lemma 3.1 Write $U_j = \sum_{\ell > j} R_\ell$. Then $\tilde R^*(z) = \sum_{j=0}^{k-1} U_j(z^j - 1)$ and so

$$\begin{aligned} \|\tilde R^*(z)\| &\ll |z-1| \sum_{j=0}^{M} j\|U_j\| e^{ja} + \sum_{j=M}^{k} \|U_j\| e^{ja} \\ &\ll |z-1| M^{1-q} \sum_{j=0}^{M} j^q \|U_j\| + M^{-q} \sum_{j=M}^{k} j^q \|U_j\|. \end{aligned}$$

Now take $M \approx 1/|z-1|$ to deduce that $\|\tilde R^*\| \ll |z-1|^q S_q(k,a)$. Hence the result follows from Corollary 3.5. ∎

4 Analyticity of $L^*(z)$ and the main result

In this section, we show how to pass from $T^*(z) = \sum_{n=0}^{\infty} 1_Y L^{*n} 1_Y z^n : \mathcal{B}(Y) \to \mathcal{B}(Y)$ to $L^*(z) = \sum_{n=0}^{\infty} L^{*n} z^n : \mathcal{B}(\Delta^*) \to L^1(\Delta^*)$. We continue to suppose that F is a good inducing scheme. (Recall that $\mathcal{B}(\Delta^*)$ is the collection of exchangeable observables.)

If in addition $\mathcal{B}(Y)$ is embedded in $L^\infty(Y)$, then our results are identical to those for $T^*(z)$ while in general we have to be content with cruder estimates that are still sufficient for the results mentioned in the introduction.

Let P_{Δ^*} denote the projection $P_{\Delta^*} v = \int_{\Delta^*} v \, d\mu_{\Delta^*}$.

Lemma 4.1 *(i) Suppose that F is a good inducing scheme and in addition that $\mathcal{B}(Y)$ is embedded in $L^\infty(Y)$. Let $a = a(k)$ be such that $\lim_{k\to\infty} a^r S_r(k,a) = 0$ for some $r \in (0,1]$. Then there exists $k_0 \ge 1$ such that for any $q \in (0,1]$, $k \ge k_0$,*

$$L^*(z) = (1-z)^{-1} P_{\Delta^*} + H^*(z) + E^*(z), \tag{4.1}$$

where $E^*(z)$ is a polynomial of degree at most $k-1$, $H^*(z)$ is analytic on the disk $\mathbb{D}_a$ and $\sup_{z\in\mathbb{D}_a} |z-1|^{1-q}\|H^*(z)\| \ll S_q(k,a)$.

(ii) In the general case of good inducing schemes, the same result holds except that $\sup_{z\in\mathbb{D}_a} \|H^*(z)\| \ll k^2 e^{2ka}$.

We can now state and prove our main result.

Theorem 4.2 *(i) Suppose that F is a good inducing scheme and in addition that $\mathcal{B}(Y)$ is embedded in $L^\infty(Y)$. Let $a = a(k)$ be such that $\lim_{k\to\infty} a^r S_r(k,a) = 0$ for some $r \in (0,1]$. Let $q \in (0,1]$. Then there exists $C > 0$, $k_0 \ge 1$ such that*

$$|\rho_{v,w}(n)| \le C|v|_\infty |w|_\infty \Big(\sum_{j>k} \mu(\varphi > j) + n\mu(\varphi > k)\Big) + C\|v\||w|_\infty S_q(k,a)e^{-na},$$

for all $v \in \mathcal{B}(X)$, $w \in L^\infty(X)$, $n \ge k \ge k_0$.

(ii) In the general case of good inducing schemes, the same result holds but with $S_q(k,a)$ replaced by $k^2 e^{2ka}$.

Proof Suppose that we are in case (i). Write $H^*(z) = \sum_{n=0}^\infty H_n^* z^n$, $E^*(z) = \sum_{j=0}^{k-1} E_n^* z^n$. Equating coefficients in (4.1) on the open unit disk $\mathbb{D}$, we obtain

$$L^{*n} = P_{\Delta^*} + H_n^* + E_n^*,$$

for $n \ge 1$, $k \ge k_0$. We claim that $\|H_n^*\| \ll S_q(k,a)e^{-na}$ for all $n \ge 1$, $k \ge k_0$. It follows that $\|L_n^* - P_{\Delta^*}\| = \|H_n^*\| \ll S_q(k,a)e^{-na}$ for all $n \ge k \ge k_0$. Hence $|\rho^*_{v,w}(n)| \ll \|v\||w|_1 S_q(k,a)e^{-na}$ for all $v \in \mathcal{B}(\Delta)$, $w \in L^1(\Delta)$. The result follows from this estimate combined with (2.1).

It remains to prove the claim. Since H^* is analytic on $\mathbb{D}_a$, $\|H_n^*\| \ll \int_\Gamma \|H^*(z)z^{-n}\| dz$ where Γ is the boundary circle of $\mathbb{D}_a$ (for a slightly smaller a). Hence

$$\|H_n^*\| \ll e^{-na} \int_0^{2\pi} \|H^*(e^{a+ib})\|\, db \ll S_q(k,a)e^{-na} \int_0^{2\pi} |e^{a+ib} - 1|^{-(1-q)}\, db.$$

But $|e^{a+ib} - 1| \ge |e^a \sin b| \ge |\sin b|$, so

$$\begin{aligned}\int_0^{2\pi} |e^{a+ib} - 1|^{-(1-q)}\, db &\le \int_0^{2\pi} |\sin b|^{-(1-q)}\, db = 4\int_0^{\pi/2} |\sin b|^{-(1-q)}\, db \\ &\ll \int_0^{\pi/2} |b|^{-(1-q)}\, db \ll 1,\end{aligned}$$

completing the proof of the claim and hence of case (i). The proof of case (ii) is similar. ∎

Remark 4.3 The statement of Theorem 4.2(i) is sufficiently general for all of our applications except in Example 5.5 where it is necessary to improve the factor $S_q(k,a)$. Such improvements can be achieved by modifying the estimate of $\tilde{R}(z)$ obtained at the end of the proof of Lemma 3.1.

In the remainder of this section, we prove Lemma 4.1. We focus on case (i), sketching the differences for case (ii) at the end of the proof.

Write $(L^{*n}v)(x) = \sum_{f^{*n}u=x} g_n^*(u)v(u)$. Following Gouëzel [6], define operator-valued polynomials

$$A^*(z) : L^\infty(Y) \to L^1(\Delta), \quad D^*(z) : \mathcal{B}(\Delta^*) \to \mathcal{B}(Y), \quad E^*(z) : L^\infty(\Delta^*) \to L^1(\Delta^*).$$

as follows:

$$A^*(z) = \sum_{n=0}^{k-1} A_n^* z^n, \quad (A_n^* v)(x) = \sum_{\substack{f^{*n}y=x \\ y\in Y;\ f^*y\notin Y,\dots,f^{*n}y\notin Y}} g_n^*(y)v(y),$$

$$D^*(z) = \sum_{n=0}^{k-1} D_n^* z^n, \quad (D_n^* v)(y) = \sum_{\substack{f^{*n}u=y \\ u\notin Y,\dots,f^{*n-1}u\notin Y;\ f^{*n}u\in Y}} g_n^*(u)v(u),$$

$$E^*(z) = \sum_{n=1}^{k-1} E_n^* z^n, \quad (E_n^* v)(x) = \sum_{\substack{f^{*n}u=x \\ u\notin Y,\dots,f^{*n}u\notin Y}} g_n^*(u)v(u).$$

(We adopt the convention that $(A_0^*v)(y,\ell) = v(y)$ for $\ell = 0$ and is zero otherwise, and that $(D_0^*v)(y) = v(y,0)$.)

As in [6], we observe that

$$L^{*n} = \textstyle\sum_{n_1+n_2+n_3=n} A_{n_1}^* T_{n_2}^* D_{n_3}^* + E_n^*.$$

Hence $L^*(z) = A^*(z)T^*(z)D^*(z) + E^*(z) : \mathcal{B}(\Delta^*) \to L^1(\Delta^*)$.

Proposition 4.4 *Let $k \ge 1$, $0 \le n \le k-1$. Then*

(a) For $v \in L^1(\Delta^*)$, $(A_n^*v)(y,\ell) = v(y)$ if $n = \ell$ and is zero otherwise.

(b) $(A^*(1)v)(y,\ell) = v(y)$ for all $(y,\ell) \in \Delta^*$.

(c) For all $v \in \mathcal{B}(\Delta^*)$, $D_n^*v = R(1_{\{\varphi>n\}}\hat v_{k,n}) = \sum_{j>n} R_j \hat v_{k,n}$, where $\hat v_{k,n} \in \mathcal{B}(Y)$ is derived from v (so in particular, $\|\hat v_{k,n}\| \le \|v\|$ for all k,n).

(d) $D^*(1)v = RV^*$ where $V^*(y) = \sum_{\ell=0}^{\varphi^*(y)-1} v(y,\ell)$.

Proof Parts (a) and (b) are immediate from the definitions.

Write $(Rv)(y) = \sum_{Fu=y} G(u)v(u)$. Then $(D_n^*v)(y) = \sum^* G(u)v(u,\varphi^*(u)-n)$ where the summation is over $u \in Y$ with $Fu = y$ and $\varphi^*(u) > n$. Parts (c) and (d) follow easily. ∎

Define $A_1^*(z) = (1-z)^{-1}(A^*(z) - A^*(1))$, $D_1^*(z) = (1-z)^{-1}(D^*(z) - D^*(1))$.

Corollary 4.5 *(a) $(1/\bar\varphi^*)A^*(1)PD^*(1) = P_{\Delta^*}$.*

(b) $\|A^*(z)\|_{L^\infty(Y)\to L^1(\Delta)} \ll 1 + a^r S_r(k,a)$ *and* $\|A_1^*(z)\|_{L^\infty(Y)\to L^1(\Delta)} \ll |z-1|^{-(1-q)} S_q(k,a)$ *for all* $q,r\in(0,1]$, $a>0$, $z\in\mathbb{D}_a$.

(c) $\|D^*(z)\|_{B(\Delta^*)\to B(Y)} \ll 1 + a^r S_r(k,a)$ *and* $\|D_1^*(z)\|_{B(\Delta^*)\to B(Y)} \ll |z-1|^{-(1-q)} S_q(k,a)$ *for all* $q,r\in(0,1]$, $a>0$, $z\in\mathbb{D}_a$.

Proof By Proposition 4.4(b,d) and the definition of μ_{Δ^*},

$$\begin{aligned}(A^*(1)PD^*(1)v)(y,\ell) = (PD^*(1)v)(y) &= \int_Y RV^*\,d\mu = \int_Y V^*\,d\mu \\ &= \int_Y \sum_{\ell=0}^{\varphi^*(y)-1} v(y,\ell)\,d\mu = \bar\varphi^* \int_{\Delta^*} v\,d\mu_{\Delta^*},\end{aligned}$$

proving part (a).

By Proposition 4.4(a), the support of A_n^* has measure $\mu(\varphi\ge n)$ and $|A_n^* v|_\infty \le |v|_\infty$. It follows that $|A_n^* v|_1 \le \mu(\varphi\ge n)|v|_\infty$. In other words, $\|A_n^*\| \le \mu(\varphi\ge n)$. Hence the estimates in part (b) are obtained in exactly the same way as the estimates for $R_1^*(z)$ and $\tilde R^*(z)$ in the proof of Lemma 3.1.

By Proposition 4.4(c), $\|D_n^*\| \ll \sum_{j>n}\|R_j^*\|$. Hence the estimates in part (c) are again obtained in exactly the same way as the estimates for $R_1^*(z)$ and $\tilde R^*(z)$. ∎

Proof of Lemma 4.1 By Lemma 3.1 and Corollary 4.5(a),

$$\begin{aligned}A^*(z)T^*(z)D^*(z) = (1-z)^{-1}(1/\bar\varphi^*)\Big\{&A^*(1)PD^*(1) + (A^*(z)-A^*(1))PD^*(1) \\ &+ A^*(z)P(D^*(z)-D^*(1))\Big\} \\ &+ A^*(z)(T^*(z)-(1-z)^{-1}(1/\bar\varphi^*)P)D^*(z)\end{aligned}$$

$$= (1-z)^{-1}P_{\Delta^*} + (1/\bar\varphi^*)A_1^*(z)PD^*(1) + (1/\bar\varphi^*)A^*(z)PD_1^*(z) + A^*(z)J^*(z)D^*(z),$$

and so

$$L^*(z) = (1-z)^{-1}P_{\Delta^*} + (1/\bar\varphi^*)A_1^*(z)PD^*(1) + (1/\bar\varphi^*)A^*(z)PD_1^*(z) + A^*(z)J^*(z)D^*(z) + E^*(z).$$

Hence, case (i) follows immediately from Lemma 3.1 and Corollary 4.5(b,c).

In case (ii), we replace Corollary 4.5(b) by the crude estimates $\|A^*(z)\|_{L^1(Y)\to L^1(\Delta)} \ll ke^{ka}$ and $\|A_1^*(z)\|_{L^1(Y)\to L^1(\Delta)} \ll k^2e^{2ka}$. ∎

5 Examples

In this section, we consider a number of special cases of Theorem 4.2, including the proofs of the results stated in the introduction. In Subsection 5.3, we verify that the AFN maps described in the introduction have the desired properties.

5.1 Calculations for good inducing schemes

In this subsection, we describe results that follow from Theorem 4.2(ii). It is assumed that $F : Y \to Y$ is a good inducing scheme with $\mathcal{B}(Y)$ embedded in $L^1(Y)$, and that $\mathcal{B}(X)$ is exchangeable. The only control required on $\|R_n\|$ is that $\sum_{n=1}^{\infty} n^{\epsilon} \sum_{j>n} \|R_j\| < \infty$ for some $\epsilon > 0$.

Proof of Theorem 1.11 Let $a(k) = \frac{1}{2}\log k/k$. For $r < \epsilon$, we compute that

$$S_r(k,a) \le S_r(k,0)e^{ka} \ll e^{ka} = k^{\frac{1}{2}},$$

and so $a^r S_r(k,a) \to 0$. By Theorem 4.2(ii), we obtain the estimate

$$|\rho(n)| \ll \sum_{j>k} \mu(\varphi > j) + n\mu(\varphi > k) + O(k^3 e^{-\frac{1}{2}n\log k/k}).$$

Now take $k = \delta n$ with $\delta = 1/(2p+6)$. ∎

Example 5.1 (Polynomially decreasing sequences [6, Définition 2.2.11]) Suppose that $\sum_{j>n} \|R_j\| = O(1/n^{1+\epsilon})$ for some $\epsilon > 0$, and that $\mu(\varphi > n) \ll u_n$ where u_n has the property that there exists a constant $C > 0$ such that $u_j \le Cu_n$ for all $n \ge 1$, $j \ge n/2$. Then we obtain the optimal upper bound $\rho(n) \ll \sum_{j>n} u_j$.

To see this, first observe that $u_j \le C^k u_n$ for all $j \ge n/2^k$ and taking $2^k \approx n$ we deduce that $n^{-p} \ll u_n$ for some $p > 0$. Also, for any $\delta > 0$ there exists $C = C(\delta)$ such that $u_{[\delta n]} \le Cu_n$. It follows that $\sum_{j>\delta n} \mu(\varphi > j) \ll \sum_{j>n} u_j$.

Finally, $\frac{n}{2}\mu(\varphi > n) \le \sum_{j\ge n/2} \mu(\varphi > j) \ll \sum_{j>n/2} u_j \ll \sum_{j>n} u_{j/2} \ll \sum_{j>n} u_j$ so $n\mu(\varphi > \delta n) \ll \sum_{j>\delta n} u_j \ll \sum_{j>n} u_j$. This accounts for all the terms in Theorem 1.11.

Example 5.2 (Regularly varying sequences, $\beta > 0$) We continue to assume that $\sum_{j>n} \|R_j\| = O(1/n^{1+\epsilon})$ for some $\epsilon > 0$. Suppose further that $\mu(\varphi > n) \ll u_n = \ell(n)/n^{\beta+1}$ where $\beta > 0$ and ℓ is a slowly varying function. (Recall that $\ell : (0,\infty) \to (0,\infty)$ is *slowly varying* if $\lim_{x\to\infty} \ell(\lambda x)/\ell(x) = 1$ for all $\lambda > 0$.)

Regularly varying sequences are clearly polynomially decreasing, so we can apply the result of Example 5.1. Moreover, $\sum_{j>n} u_j \ll \ell(n)/n^{\beta}$ by a result of Karamata (see [4, Theorem 1, p. 273]). Hence we obtain the optimal upper bound $|\rho(n)| \ll \ell(n)/n^{\beta}$.

Finally, we consider the standard case of exponential decay of correlations.

Example 5.3 (Exponential decay rates) If $\|R_n\| = O(e^{-cn})$, $c > 0$, then we obtain exponential decay of correlations as expected. A slight reformulation of Theorem 4.2(ii) is required, where we modify the condition $\lim_{k\to\infty} a^r S_r(k,a) = 0$. Indeed the only places where the condition is used (rather than simply boundedness) is in Step 3 of the proof of Proposition 3.2 and in ensuring that $(1/\varphi^*)|\tilde{\lambda}^*| < \frac{1}{2}$ in the proof of Corollary 3.5. For these it suffices that for any $\epsilon > 0$, there exists $a = a(k)$ and $r \in (0, 1]$ such that $a^r S_r(k, a) < \epsilon$.

Under the assumption $\|R_n\| = O(e^{-cn})$, this new condition can be satisfied with $r = 1$ and a chosen to be a sufficiently small constant $a \equiv \epsilon_1 \in (0, c)$. (Taking $\epsilon_1 < c$ ensures that $S_1(k, a)$ is bounded; the other requirements on ϵ_1 are less explicit.) Let $n = k$. Since $\mu(\varphi = n) \ll \|R_n\| = O(e^{-cn})$, we obtain $|\rho(n)| \ll n^2 e^{-\epsilon_1 n}$.

5.2 Calculations for inducing schemes with $\mathcal{B}(Y) \subset L^\infty(Y)$

In this subsection, we suppose that $F : Y \to Y$ is an excellent inducing scheme and that $\mathcal{B}(Y)$ is embedded in $L^\infty(X)$. As usual, we suppose that $\mathcal{B}(X)$ is exchangeable.

Since $\mathcal{B}(Y)$ is embedded in $L^\infty(X)$, we can appeal to part (i) of Theorem 4.2. Since F is excellent, hypotheses on φ are inherited by $\|R_n\|$. (The results in this subsection can be formulated for good inducing schemes by imposing conditions on $\|R_n\|$ directly but the ensuing results are suboptimal.)

Proof of Theorem 1.4 We take $q = r = 1$ in Theorem 4.2(i). Let $a = \frac{1}{2}k^{-1}\log\log k$. Then, $S_1(k, a) \le \sum_{j=1}^{k}(\log^{-1} j)e^{ak} \ll k \log^{-\frac{1}{2}} k$, and so $aS_1(k, a) \to 0$. Moreover, $S_1(k, a)e^{-na} \ll k(\log^{-\frac{1}{2}} k)e^{-na} \to 0$ with $n = 2k\log k/\log\log k$. In addition, $n\mu(\varphi > k) \ll (\log\log k)^{-1} \to 0$. Finally, $\sum_{j>k}\mu(\varphi > j) \to 0$ since $\varphi \in L^1(Y)$. ∎

Example 5.4 We give an example where $\varphi \in L^1$ and $\mu(\varphi > n) = O(1/(n\log n))$, so that we obtain a uniform rate of decay of correlations by Theorem 1.4. However, $\mu(\varphi > n)n\log n \not\to 0$. As far as we know, this means that the example is intractable by previous approaches.

Suppose that φ takes only the values $[e^{j^2}]$, $j \ge 1$, and that $\mu(\varphi = [e^{j^2}]) = C/(j^2 e^{j^2})$ where C is chosen so that μ is a probability measure. Clearly $\varphi \in L^1$ since $\sum_{n=1}^{\infty} n\mu(\varphi = n) = \sum_{j=1}^{\infty}[e^{j^2}]C/(j^2 e^{j^2}) \le C\sum_{j=1}^{\infty} 1/j^2 < \infty$.

If $n = [e^{j^2}]$, then $\mu(\varphi > n) = \sum_{k>j} C/(k^2 e^{k^2}) \approx 1/(j^2 e^{j^2}) = 1/(e^{j^2}\log(e^{j^2})) \approx 1/(n\log n)$. Hence, $\mu(\varphi > n) = O(1/(n\log n))$ and $\mu(\varphi > n)n\log n \not\to 0$ as required.

The proof of Theorem 1.4 gives an explicit decay rate as follows. We have $\rho(n) \ll \sum_{j>k}\mu(\varphi > j) + O(1/\log\log k)$ where $n = 2k\log k/\log\log k$. Moreover,

$$\sum_{j>k}\mu(\varphi > j) \le \sum_{j>k} j\mu(\varphi = j) = \sum_{[e^{j^2}]>k}[e^{j^2}]C/(j^2 e^{j^2}) \le C\sum_{j>\log^{\frac{1}{2}} k} 1/j^2 = O(\log^{-\frac{1}{2}} k).$$

Hence $\rho(n) = O(1/\log\log k) = O(1/\log\log n)$.

Example 5.5 (Regularly varying sequences, $\beta = 0$) We consider the case of regularly varying sequences $\mu(\varphi > n) \ll \ell(n)/n$ where $\ell(n) \to 0$ as $n \to \infty$ (supposing as always that $\varphi \in L^1(Y)$). Many such examples were considered by Holland [8].

We suppose also that $\ell(n)$ is decreasing. It follows that $\ell(n)\log n$ is bounded (since $\ell(n)\log n \ll \ell(n)\sum_{j=1}^{n} 1/j \le \sum_{j=1}^{n}\ell(j)/j \ll 1$).

Take $a(k) = \frac{1}{2}k^{-1}\log(1/\ell(k))$. By Karamata,

$$S_1(k,a) \le S_1(k,0)e^{ka} \le \ell(k)^{-\frac{1}{2}}\sum_{j=1}^{k}\ell(j) \ll k\ell(k)^{\frac{1}{2}},$$

and it follows that $\lim_{k\to\infty} aS_1(k,a) = 0$.

As mentioned in Remark 4.3, we require a refinement to the estimate of $\|\tilde{R}^*\|$ at the end of the proof of Lemma 3.1. Recall that $\tilde{R}^*(z) = \sum_{j=0}^{k-1} U_j(z^j - 1)$ where $U_j = \sum_{\ell>j} R_\ell$. By assumption, $\|U_j\| \ll \ell(j)/j$. By Karamata and the assumption that $\ell(n)$ is decreasing,

$$\begin{aligned}\|\tilde{R}^*(z)\| &\le \Big\{|z-1|\sum_{j=1}^{M}\ell(j) + \sum_{j=M}^{k} j^{-1}\ell(j)\Big\}e^{ka} \le \{|z-1|M\ell(M) + \ell(M)\log k\}/\ell(k)^{\frac{1}{2}} \\ &\le \{|z-1|M\ell(M) + \ell(M)\}/\ell(k)^{\frac{3}{2}}\end{aligned}$$

so taking $M \approx 1/(z-1)$ yields $\|\tilde{R}^*(z)\| \ll \ell(1/|z-1|)\ell(k)^{-\frac{3}{2}}$ on $\mathbb{D}_a$.

Since $\ell(n)/n$ is summable, it follows that $\ell(1/\theta)(1/\theta)$ is integrable. Hence we can argue as in the proof of Theorem 4.2 to deduce that $|\rho^*(n)| \ll \ell(k)^{-\frac{3}{2}}e^{-na}$ and so

$$|\rho(n)| \ll \sum_{j>k}\ell(j)/j + n\ell(k)/k + \ell(k)^{-\frac{3}{2}}e^{-\frac{1}{2}nk^{-1}\log(1/\ell(k))}.$$

Define $\tilde{\ell}(n) = \sum_{j=n}^{\infty} j^{-1}\ell(j)$. By Karamata, $\ell(n) = o(\tilde{\ell}(n))$. Taking $n = 5k$, we obtain the upper bound $|\rho(n)| \ll \tilde{\ell}(n)$.

Example 5.6 (Stretched exponential sequences) We consider the case $\mu(\varphi > n) \ll e^{-cn^\gamma}$, where $\gamma \in (0,1)$ and $c > 0$.

Take $a = k^{-1}(ck^\gamma - (1+\epsilon)\log k)$. Since $a(k)$ is eventually decreasing, we can replace the $e^{ja(k)}$ factor in $S_q(k,a)$ by $e^{ja(j)}$. Then a calculation shows that $S_q(k,a) \ll 1 + k^{q-\epsilon}$ for all $q \in (0,1)$. In particular, $a^r S_r(k,a) \to 0$ for $r \in (0,\epsilon/\gamma)$. Taking $n = k$ we obtain $|\rho(n)| \ll n^{1+\epsilon}e^{-cn^\gamma}$ for any $\epsilon > 0$.

Even in the special setting of Young towers, this is stronger than estimates obtained by coupling [19] or cones [12]. However for Young towers, Gouëzel [6] obtains the optimal estimate $n^{1-\gamma}e^{-cn^\gamma}$. In a future paper, we show how to recover Gouëzel's result by elementary arguments.

The method described here works more generally for the case $\mu(\varphi > n) \ll e^{-g(n)}$ where $g(n)$ is an increasing sequence satisfying $g(n) = O(n^{1-\epsilon})$ for some $\epsilon > 0$ and such that $a(k) = k^{-1}(g(k) - (1+\epsilon)\log k)$ is eventually decreasing for some $\epsilon > 0$. Then we obtain $|\rho(n)| \ll n^{1+\epsilon}e^{-g(n)}$ for any $\epsilon > 0$.

5.3 AFN maps

As mentioned in the introduction, the AFN maps studied by [20] have an excellent inducing scheme with standard function space being the space $\mathrm{BV}(Y)$ of observables of bounded variations. Unfortunately, the corresponding space $\mathrm{BV}(X)$ is not exchangeable.

Instead we take $\mathcal{B}(Y)$ to be the space of piecewise bounded variation functions $v : Y \to \mathbb{R}$ with norm $\|v\| = \sup_{a\in\alpha} \|1_a v\|_{\mathrm{BV}}$. Let $\mathcal{B}(X)$ be the space of piecewise bounded variation functions $v : X \to \mathbb{R}$ with norm $\|v\| = \sup_{a\in\alpha, 0\le\ell\le\phi(a)-1} \|1_a\, v \circ T^\ell\|_{\mathrm{BV}}$. Since T^ℓ restricted to a is a homeomorphism for $0 \le \ell \le \phi(a) - 1$, it is immediate that $\mathcal{B}(X)$ is exchangeable.

It remains to show that $F : Y \to Y$ is an excellent inducing scheme relative to $\mathcal{B}(Y)$. The details are standard, so we sketch the argument. Let $\hat{R}$ denote the transfer operator with respect to Lebesgue measure. Then it follows from [16] and [20, Appendix] that $\hat{R} : \mathrm{BV}(Y) \to \mathrm{BV}(Y)$ is bounded and quasicompact, so there exist constants $C > 0$, $\tau \in (0, 1)$ such that $\|\hat{R}^n v\|_{\mathrm{BV}} \le C\tau^n \|v\|_{\mathrm{BV}}$ for all $n \ge 1$ and all $v \in \mathrm{BV}(Y)$ with $\int_Y v\, d\mu = 0$. Moreover, $d\mu = h\, dy$ where the density $h \in L^1(Y)$ satisfies $h, h^{-1} \in \mathrm{BV}(Y)$. Hence $R = h^{-1}\hat{R}h$ inherits the quasicompactness on $\mathrm{BV}(Y)$ verifying (H2)(i). Following [1] (see for example [14, Subsection 11.3]), it is possible to extend this analysis to $R(z)$ for all $z \in \bar{\mathbb{D}}$ and to verify (H2)(ii). Further, F has good distortion properties, so (H1) is easily verified. Hence F is excellent relative to $\mathrm{BV}(Y)$.

To prove excellence relative to $\mathcal{B}(Y)$, we note that $\hat{R} : \mathcal{B}(Y) \to \mathrm{BV}(Y)$ is bounded and hence defines a bounded operator on $\mathcal{B}(Y)$. (This is identical to the argument for $\mathrm{BV}(Y)$ since F satisfies a strong Rychlik condition [1, Condition (R), page 53].) Hence it is immediate from the results on $\mathrm{BV}(Y)$ that there exist constants $C > 0$, $\tau \in (0, 1)$ such that $\|R^n v\| \le C\tau^n \|v\|$ for all $n \ge 1$ and all $v \in \mathcal{B}(Y)$ with $\int_Y v\, d\mu = 0$, verifying (H2)(i). The other properties are inherited from $\mathrm{BV}(Y)$ similarly.

A Details for the truncation error

In this appendix, we give the details for the truncation error (2.1). A similar result was proved in [13] in a slightly more complicated situation. We give the details mainly for completeness and also because we obtain a slightly improved formula (though the improvement is never used).

In particular, we use a slightly better splitting for Δ, namely $\Delta = \Delta^* \dot{\cup} \Delta_{\mathrm{trunc}}$ where $\Delta_{\mathrm{trunc}} = \{(y, \ell) \in \Delta : \ell > k\}$.

Proposition A.1 *(i)* $\bar{\varphi} - \bar{\varphi}^* = \sum_{j>k} \mu(\varphi > j)$.

(ii) $\mu_\Delta(\Delta_{\mathrm{trunc}}) = (1/\bar{\varphi}) \sum_{j>k} \mu(\varphi > j)$.

Proof This is a standard computation. ∎

Proposition A.2 *For $n \geq 1$, define*

$$E_n = \{x \in \Delta^* : f^j x \in \Delta_{\text{trunc}} \textit{ for at least one } j \in \{1, \ldots, n\}\}.$$

Then $\mu_\Delta(E_n) \leq n\mu(\varphi > k)$.

Proof Write E_n as the disjoint union $E_n = \bigcup_{j=1}^n G_j$ where

$$G_j = \{f^i x \in \Delta^* \text{ for } i \in \{0, 1, \ldots, j-1\} \text{ and } f^j x \in \Delta_{\text{trunc}}\}.$$

It follows from the definition that if $x \in G_j$, then $f^j x \in \Delta_{k+1}$ where $\Delta_{k+1} = \{(y, k+1) : \varphi(y) > k\}$(the $(k+1)$'th level of the tower). Hence $\mu_\Delta(G_j) \leq \mu_\Delta(f^{-j}(\Delta_{k+1})) = \mu_\Delta(\Delta_{k+1}) = (1/\bar{\varphi})\mu(\varphi > k)$. ∎

Corollary A.3 *Suppose that $v, w : \Delta \to \mathbb{R}$ lie in L^∞. Then for all $k, n \geq 1$,*

$$|\rho(n) - \rho^*(n)| \leq C|v|_\infty |w|_\infty \{\textstyle\sum_{n>j} \mu(\varphi > j) \, + \, n\mu(\varphi > k)\}.$$

Proof First we estimate $S = \int_\Delta v\, w \circ f^n \, d\mu_\Delta - \int_{\Delta^*} v\, w \circ f^{*n} \, d\mu_{\Delta^*}$. Write

$$\begin{aligned} S &= \textstyle\int_{\Delta_{\text{trunc}}} v\, w \circ f^n \, d\mu_\Delta + \Big(\int_{\Delta^*} v\, w \circ f^n \, d\mu_\Delta - \int_{\Delta^*} v\, w \circ f^{*n} \, d\mu_\Delta \Big) \\ &\qquad + \Big(\textstyle\int_{\Delta^*} v\, w \circ f^{*n} \, d\mu_\Delta - \int_{\Delta^*} v\, w \circ f^{*n} \, d\mu_{\Delta^*} \Big) \\ &= I + II + III. \end{aligned}$$

Now $|I| \leq |v|_\infty |w|_\infty \mu_\Delta(\Delta_{\text{trunc}})$. Note that $\mu_\Delta|\Delta^* = (\bar{\varphi}^*/\bar{\varphi})\mu_{\Delta^*}|\Delta^*$ and so $III = (\bar{\varphi}^*/\bar{\varphi} - 1) \int_{\Delta^*} v\, w \circ f^n \, d\mu_{\Delta^*}$. Hence $|III| \leq (1/\bar{\varphi})(\bar{\varphi} - \bar{\varphi}^*)|v|_\infty |w|_\infty$. Next, $|II| \leq 2|v|_\infty |w|_\infty \mu_\Delta(\Delta^* \cap \{f^n \neq f^{*n}\}) \leq 2|v|_\infty |w|_\infty \mu_\Delta(E_n)$. Combining these, we obtain

$$\begin{aligned} |S| &\leq |v|_\infty |w|_\infty \{\mu_\Delta(\Delta_{\text{trunc}}) + |\bar{\varphi} - \bar{\varphi}^*| + 2\mu_\Delta(E_n)\} \\ &\leq C|v|_\infty |w|_\infty \{\sum_{n>j} \mu(\varphi > j) \, + \, n\mu(\varphi > k)\} \end{aligned}$$

by Propositions A.1 and A.2.

A similar (but simpler) calculation shows that

$$\Big|\textstyle\int_\Delta v \, d\mu_\Delta \int_\Delta w \, d\mu_\Delta - \int_{\Delta^*} v \, d\mu_{\Delta^*} \int_{\Delta^*} w \, d\mu_{\Delta^*}\Big| \leq C|v|_\infty |w|_\infty \sum_{j>k} \mu(\varphi > j),$$

and the result follows. ∎

B Nonuniformly hyperbolic systems

In this appendix, we show how our main results for nonuniformly expanding maps extend to nonuniformly hyperbolic maps modelled by Young towers [18, 19]. Even in the case of polynomial tails, this result has been missing from the literature. (In the case of exponential tails, Young [18] explicitly considers both the nonuniformly expanding and nonuniformly hyperbolic situations, but the subexponential tail paper [19] is set entirely in the nonuniformly expanding framework.)

A method for passing from nonuniformly expanding maps to nonuniformly hyperbolic systems with subexponential tails was shown to one of us by Sébastien Gouëzel [7] based on ideas in [3]. Here, we combine these ideas with dynamical truncation.

Let $T : M \to M$ be a diffeomorphism (possibly with singularities) defined on a Riemannian manifold (M, d). Fix a subset $Y \subset M$. It is assumed that there is a "product structure": namely a family of "stable disks" $\{W^s\}$ that are disjoint and cover Y, and a family of "unstable disks" $\{W^u\}$ that are disjoint and cover Y. Each stable disk intersects each unstable disk in precisely one point. The stable and unstable disks containing y are labelled $W^s(y)$ and $W^u(y)$.

(P1) There is a partition $\{Y_j\}$ of Y and integers $\varphi_j \geq 1$ such that $T^{\varphi_j}(W^s(y)) \subset W^s(T^{\varphi_j}y)$ for all $y \in Y_j$.

Define the return time function $\varphi : Y \to \mathbb{Z}^+$ by $\varphi|_{Y_j} = \varphi_j$ and the induced map $F : Y \to Y$ by $F(y) = T^{\varphi(y)}(y)$.

Let s denote the *separation time* with respect to the map $F : Y \to Y$. That is, if $y, z \in Y$, then $s(y, z)$ is the least integer $n \geq 0$ such that $F^n x$, $F^n y$ lie in distinct partition elements of Y.

(P2) There exist constants $C \geq 1$, $\gamma_0 \in (0, 1)$ such that

- (i) If $z \in W^s(y)$, then $d(F^n y, F^n z) \leq C\gamma_0^n$,
- (ii) If $z \in W^u(y)$, then $d(F^n y, F^n z) \leq C\gamma_0^{s(y,z)-n}$,
- (iii) If $y, z \in Y$, then $d(T^j y, T^j z) \leq Cd(y, z)$ for all $0 \leq j < \min\{\varphi(y), \varphi(z)\}$.

Let $\bar{Y} = Y/\sim$ where $y \sim z$ if $y \in W^s(z)$ and define the partition $\{\bar{Y}_j\}$ of $\bar{Y}$. We obtain a well-defined return time function $\varphi : \bar{Y} \to \mathbb{Z}^+$ and induced map $\bar{F} : \bar{Y} \to \bar{Y}$.

(P3) The map $\bar{F} : \bar{Y} \to \bar{Y}$ and partition $\{\bar{Y}_j\}$ separate points in $\bar{Y}$. (It follows that $d_\theta(y, z) = \theta^{s(y,z)}$ defines a metric on $\bar{Y}$ for each $\theta \in (0, 1)$.)

(P4) There exists an invariant ergodic probability measure $\mu_{\bar{Y}}$ on $\bar{Y}$ such that $F : \bar{Y} \to \bar{Y}$ is a Gibbs-Markov map in the sense of Example 1.7 and $\varphi : \bar{Y} \to \mathbb{Z}^+$ is integrable.

From (P4), a standard construction leads to an invariant probability measure μ_Y on Y such that $\bar{\pi}_*\mu_Y = \mu_{\bar{Y}}$ where $\bar{\pi} : Y \to \bar{Y}$ is the quotient map. There is also a standard method to pass from μ_Y to a measure ν on M which we recall now. As in Section 2.1, starting from $\bar{F} : \bar{Y} \to \bar{Y}$ and $\varphi : \bar{Y} \to \mathbb{Z}^+$, we can form a *quotient tower* $\bar{\Delta}$ and a *quotient tower map* $\bar{f} : \bar{\Delta} \to \bar{\Delta}$ such that $\bar{F} = \bar{f}^{\varphi} : \bar{Y} \to \bar{Y}$ is a *first* return map for $\bar{f}$. Then $\mu_{\bar{\Delta}} = (\mu_{\bar{Y}} \times \text{counting})/\int_{\bar{Y}} \varphi \, d\mu_{\bar{Y}}$ is an $\bar{f}$-invariant probability measure on $\bar{\Delta}$.

Similarly, starting from $F : Y \to Y$ and $\varphi : Y \to \mathbb{Z}^+$, we can form a tower Δ and tower map $f : \Delta \to \Delta$ such that $F = f^{\varphi} : Y \to Y$ is a *first* return map for f. Again, $\mu_\Delta = (\mu \times \text{counting})/\int_Y \varphi \, d\mu_Y$ is an f-invariant probability measure on Δ. Define the semiconjugacy $\pi : \Delta \to M$, $\pi(y, \ell) = T^\ell y$. Then $\nu = \pi_*\mu_\Delta$ is the desired measure on M. (We omit the additional assumptions in Young [18] that guarantee that ν is an SRB measure. The results in this appendix do not rely on this property.)

Let $v_0, w_0 : M \to \mathbb{R}$ be C^η observables ($\eta \in (0,1)$) and define the correlation function $\rho_{v_0,w_0}(n) = \int_M v_0 \, w_0 \circ T^n \, d\nu - \int_M v_0 \, d\nu \int_M w_0 \, d\nu$. We obtain the following analogue of Theorem 4.2(i).

Theorem B.1 *Let $a = a(k)$ be such that $\lim_{k\to\infty} a^r S_r(k,a) = 0$ for some $r \in (0,1]$. Let $q \in (0,1]$. Then there exists $C > 0$, $k_0 \geq 1$ such that*

$$|\rho_{v_0,w_0}(n)| \leq C|v_0|_\infty |w_0|_\infty \Big(\sum_{j>k} \mu(\varphi > j) + n\mu(\varphi > k)\Big) + C\|v_0\|_{C^\eta} \|w_0\|_{C^\eta} S_q(k,a) e^{-\frac{1}{2}na},$$

for all $v_0, w_0 \in C^\eta(M)$, $n \geq k \geq k_0$.

Remark B.2 Thus, we obtain identical results for the nonuniformly hyperbolic case as for the nonuniformly expanding case, except that $a(k)$ is replaced by $\frac{1}{2}a(k)$. In particular, we obtain optimal results for polynomial decay, and more generally for polynomially decreasing sequences. In addition, Corollary 1.3 and Theorem 1.4 remain valid. The only result that deteriorates in passing to the nonuniformly hyperbolic setting is the estimate for stretched exponential decay in Example 5.6 where we obtain the decay rate $O(n^{1+\epsilon} e^{-\frac{1}{2}cn^\gamma})$.

In the remainder of this appendix, we prove Theorem B.1.

Decay of correlations on Δ Given C^η observables $v_0, w_0 : M \to \mathbb{R}$, let $v = v_0 \circ \pi, w = w_0 \circ \pi : \Delta \to \mathbb{R}$ be the lifted observables. Since $\pi : \Delta \to M$ is a semiconjugacy and $\nu = \pi_*\mu_\Delta$, to prove Theorem B.1 it is equivalent to estimate the correlation function $\rho_{v,w}(n) = \int_\Delta v \, w \circ f^n \, d\mu_\Delta - \int_\Delta v \, d\mu_\Delta \int_\Delta w \, d\mu_\Delta$.

Dynamical truncation For $k \geq 1$ fixed, set $\varphi^* = \min\{\varphi, k\}$ to form a truncated tower map $f^* : \Delta^* \to \Delta^*$ (with invariant probability measure μ_{Δ^*}). Let $\rho^*_{v,w}(n) = \int_{\Delta^*} v \, w \circ f^{*n} \, d\mu^* - \int_{\Delta^*} v \, d\mu^* \int_{\Delta^*} w \, d\mu^*$. We obtain the same truncation error (2.1) as in the nonuniformly hyperbolic case. Hence it remains to prove under the assumptions of Theorem B.1 that

$$|\rho^*_{v,w}(n)| \le C\|v_0\|_{C^\eta}\|w_0\|_{C^\eta} S_q(k,a)e^{-\frac{1}{2}na}. \tag{B.1}$$

Quotient towers and function spaces We use the separation time for $F : Y \to Y$ to define a separation time on Δ: define $s((y,\ell),(z,m)) = s(y,z)$ if $\ell = m$ and 0 otherwise. This drops down to separation times s on $\bar\Delta$ and $\bar Y$.

Given $\theta \in (0,1)$, we define the symbolic metric d_θ on $\bar\Delta$ by setting $d_\theta(p,q) = \theta^{s(p,q)}$. In particular, d_θ is a metric on $\bar Y$. Define the spaces $\mathcal{B}(\bar\Delta)$, $\mathcal{B}(\bar Y)$ of d_θ-Lipschitz observables on $\bar\Delta$ and $\bar Y$ respectively. Then $\mathcal{B}(\bar Y)$ satisfies our main hypotheses (H1) and (H2), and $\mathcal{B}(\bar\Delta)$ is exchangeable.

Nonuniform expansion/contraction Recall that $\pi : \Delta^* \to M$ denotes the projection $\pi(y,\ell) = T^\ell y$. For $p = (x,\ell), q = (y,\ell) \in \Delta^*$, we write $q \in W^s(p)$ if $y \in W^s(x)$ and $q \in W^u(p)$ if $y \in W^u(x)$. Conditions (P2) translate as follows.

(P2′) There exist constants $C \ge 1$, $\gamma_0 \in (0,1)$ such that for all $p, q \in \Delta^*$, $n \ge 1$,

- (i) If $q \in W^s(p)$, then $d(\pi f^{*n}p, \pi f^{*n}q) \le C\gamma_0^{\psi^*_n(p)}$, and
- (ii) If $q \in W^u(p)$, then $d(\pi f^{*n}p, \pi f^{*n}q) \le C\gamma_0^{s(p,q)-\psi^*_n(p)}$,

where $\psi^*_n(p) = \#\{j = 0,\dots,n-1 : f^{*j}p \in Y\}$ is the number of returns of p to Y by time n.

Remark B.3 These properties can be defined at the level of the nontruncated tower Δ. Since F is independent of k, the constants γ_0 and C are unchanged by truncation and hence are independent of k. Also, $s(p,q)$ is independent of k. Of course, $\psi^*_n(p)$ decreases monotonically with k, and we have the estimate $n/k \le \psi^*_n \le n$.

Proposition B.4 $d(\pi f^{*n}p, \pi f^{*n}q) \le C\gamma_0^{\min\{\psi^*_n(p),s(p,q)-\psi^*_n(p)\}}$ *for all* $p, q \in \Delta$, $n \ge 1$.

Proof This is immediate from conditions (P2′) and the product structure on Y. ∎

Approximation of observables Let $v = v_0 \circ \pi : \Delta^* \to \mathbb{R}$ be the lift of a C^η observable $v_0 : M \to \mathbb{R}$. For each $n \ge 1$, define $\tilde v_n : \Delta^* \to \mathbb{R}$,

$$\tilde v_n(p) = \inf\{v(f^{*n}q) : s(p,q) \ge 2\psi^*_n(p)\}.$$

We list some standard properties of $\tilde v_n$. Recall that L^* is the transfer operator corresponding to $\bar f^* : \bar\Delta \to \bar\Delta$.

Proposition B.5 *The function* $\tilde v_n$ *lies in* $L^\infty(\Delta^*)$ *and projects down to a Lipschitz observable* $\bar v_n : \bar\Delta^* \to \mathbb{R}$. *Moreover, setting* $\gamma = \gamma_0^\eta$ *and* $\theta = \gamma^{\frac{1}{2}}$, *(a)* $|\bar{v}_n|_\infty = |\tilde{v}_n|_\infty \le |v_0|_\infty$.

(b) $|v \circ f^{*n}(p) - \tilde{v}_n(p)|_\infty \le C\|v_0\|_{C^\eta}\gamma^{\psi_n^*(p)}$ *for* $p \in \Delta^*$.

(c) $\|L^{*n}\bar{v}_n\|_\theta \le C\|v_0\|_{C^\eta}$.

Proof If $s(p,q) \ge 2\psi_n^*(p)$, then $\tilde{v}_n(p) = \tilde{v}_n(q)$. It follows that $\tilde{v}_n$ is piecewise constant on a measurable partition of Δ^*, and hence is measurable, and that $\bar{v}_n$ is well-defined. Part (a) is immediate.

Recall that $v = v_0 \circ \pi$ where $v_0 : M \to \mathbb{R}$ is C^η. Let $p \in \Delta^*$. By Proposition B.4 and the definition of $\tilde{v}_n$,

$$\begin{aligned} |v \circ f^{*n}(p) - \tilde{v}_n(p)| &= |v_0(\pi f^{*n}p) - v_0(\pi f^{*n}q)| \le \|v_0\|_{C^\eta} d(\pi f^{*n}p, \pi f^{*n}q)^\eta \\ &\le C^*\gamma^{\min\{\psi_n^*(p), s(p,q)-\psi_n^*(p)\}}, \end{aligned}$$

where q is such that $s(p,q) \ge 2\psi_n^*(p)$. In particular, $s(p,q) - \psi_n^*(p) \ge \psi_n^*(p)$, so we obtain part (b).

To prove (c), recall that $(L^{*n}\bar{v}_n)(\bar{p}) = \sum_{\bar{f^*}^n\bar{q}=\bar{p}} g_n(\bar{q})\bar{v}_n(\bar{q})$ where g is the weight function. It is immediate that $|L^{*n}\bar{v}_n|_\infty \le |\bar{v}_n|_\infty \le |v_0|_\infty$. Write

$$\begin{aligned} (L^{*n}\bar{v}_n)(\bar{p}_1) - (L^{*n}\bar{v}_n)(\bar{p}_2) = &\sum_{\bar{f^*}^n\bar{q}_1=\bar{p}_1} g_n(\bar{q}_1)(\bar{v}_n(\bar{q}_1) - \bar{v}_n(\bar{q}_2)) \\ &+ \sum_{\bar{f^*}^n\bar{q}_1=\bar{p}_1} (g_n(\bar{q}_1) - g_n(\bar{q}_2))\bar{v}_n(\bar{q}_2). \end{aligned} \tag{B.2}$$

Naturally, we pair up preimages so that $s(\bar{q}_1, \bar{q}_2) = \psi_n^*(\bar{q}_1) + s(\bar{p}_1, \bar{p}_2)$. We then choose $q_1, q_2 \in \Delta^*$ that project onto $\bar{q}_1, \bar{q}_2 \in \bar{\Delta}^*$, so

$$s(q_1, q_2) = s(\bar{q}_1, \bar{q}_2) = \psi_n^*(\bar{q}_1) + s(\bar{p}_1, \bar{p}_2). \tag{B.3}$$

By standard arguments, the second term in (B.2) contributes $C|v_0|_\infty$ to the norm of $L^{*n}\bar{v}_n$. We claim that $|\bar{v}_n(\bar{q}_1) - \bar{v}_n(\bar{q}_2)| \le C\|v_0\|_{C^\eta}\gamma^{\frac{1}{2}s(\bar{p}_1,\bar{p}_2)}$. Taking $\theta = \gamma^{\frac{1}{2}}$, it then follows that the first term in (B.2) contributes $C\|v_0\|_{C^\eta}$ to the norm of $L^{*n}\bar{v}_n$.

It remains to verify the claim. Write

$$\bar{v}_n(\bar{q}_1) - \bar{v}_n(\bar{q}_2) = v \circ f^{*n}(\hat{q}_1) - v \circ f^{*n}(\hat{q}_2),$$

where $\hat{q}_1, \hat{q}_2 \in \Delta^*$ satisfy

$$s(\hat{q}_j, q_j) \ge 2\psi_n^*(\bar{q}_j). \tag{B.4}$$

Moreover, $\bar{v}_n(\bar{q}_1) = \bar{v}_n(\bar{q}_2)$ if $s(q_1, q_2) \ge 2\psi_n^*(\bar{q}_1)$, so we may suppose without loss that

$$s(q_1, q_2) \le 2\psi_n^*(\bar{q}_1). \tag{B.5}$$

As in part (b),

$$|v \circ f^{*n}(\hat{q}_1) - v \circ f^{*n}(\hat{q}_2)| \le C\|v_0\|_{C^\eta}\gamma^{\min\{\psi_n^*(\hat{q}_1), s(\hat{q}_1,\hat{q}_2)-\psi_n^*(\hat{q}_1)\}}. \tag{B.6}$$

By (B.3) and (B.4),

$$s(\hat{q}_1,\hat{q}_2)-\psi_n^*(\bar{q}_1) \ge \min\{s(q_1,q_2), s(\hat{q}_1,q_1), s(\hat{q}_2,q_2)\}-\psi_n^*(\bar{q}_1) \ge \min\{s(\bar{p}_1,\bar{p}_2), \psi_n^*(\bar{q}_1)\}.$$

By (B.4) and (B.5),

$$\psi_n^*(\hat{q}_1) = \psi_n^*(\bar{q}_1) \ge \tfrac{1}{2}s(\bar{q}_1,\bar{q}_2) \ge \tfrac{1}{2}s(\bar{p}_1,\bar{p}_2).$$

Substituting these into (B.6) establishes the claim. ∎

The next property draws on ideas from [3, Lemma 4.4].

Lemma B.6 *Suppose that $a = a(k)$ satisfies $\lim_{k\to\infty} aS_0(k,a) = 0$. Let $r \in (0,1]$. There exists $k_0 \ge 1$ such that*

$$|v \circ f^{*n} - \tilde{v}_n|_1 \ll (1 + a^r S_r(k,a))^2 e^{-na}\|v_0\|_{C^\eta},$$

for all $n \ge k \ge k_0$.

Proof By Proposition B.5(b), $|v \circ f^{*n}(p) - \tilde{v}_n(p)| \ll \gamma^{\psi_n^*(p)}\|v_0\|_{C^\eta}$. Note that $\psi_n^* = \sum_{j=0}^{n-1}\psi \circ f^{*j}$ where $\psi = 1_Y$. We have

$$|v \circ f^{*n} - \tilde{v}_n|_1/\|v_0\|_{C^\eta} \ll \int_{\Delta^*}\gamma^{\psi_n^*}\,d\mu^* = \int_{\bar{\Delta}^*}\gamma^{\psi_n^*}\,d\mu^* = \int_{\bar{\Delta}^*}L^{*n}\gamma^{\psi_n^*}\,d\mu^* = \int_{\bar{\Delta}^*}L_\gamma^{*n}1\,d\mu,$$

where L_γ^* is the twisted transfer operator $L_\gamma^* v = L^*(\gamma^\psi v)$.

We estimate L_γ^{*n} using truncated renewal operators. Define

$$T_{n,\gamma}^* = 1_Y L_\gamma^{*n} 1_Y, \qquad T_\gamma^*(z) = \sum_{n=0}^\infty T_{n,\gamma}^* z^n,$$

$$R_{n,\gamma}^* = 1_Y L_\gamma^{*n} 1_{\{\varphi^*=n\}}, \qquad R_\gamma^*(z) = \sum_{n=1}^\infty R_{n,\gamma}^* z^n.$$

Then the renewal equation takes the form $T_\gamma^*(z) = (I - R_\gamma^*(z))^{-1}$, for $z \in \mathbb{D}$. Throughout, $\gamma \in (0,1)$ is fixed.

Next, we observe that

$$R_{n,\gamma}^* v = L_\gamma^{*n}(1_{\{\varphi^*=n\}}v) = L^{*n}(\gamma^{\psi_n^*}1_{\{\varphi^*=n\}}v) = \gamma R_n^* v.$$

In particular, $R_\gamma^*(z) = \gamma R^*(z)$ for $z \in \mathbb{C}$. Similarly, we can define $R_\gamma(z)$ and deduce that $R_\gamma(z) = \gamma R(z)$, $z \in \bar{\mathbb{D}}$. Hence, the spectral radius of $R_\gamma(z)$ is at most γ for all

$z \in \bar{\mathbb{D}}$. It follows that $\sup_{z\in\bar{\mathbb{D}}} \|(I - R_\gamma(z))^{-1}\| < \infty$ for $z \in \bar{\mathbb{D}}$. We proceed as in the proof of Proposition 3.2 to deduce that for $k \geq k_0$, first $\sup_{z\in\bar{\mathbb{D}}} \|(I - R^*_\gamma(z))^{-1}\| < \infty$, and then that

$$\sup_{z\in\mathbb{D}_a} \|T^*_\gamma(z)\| = \sup_{z\in\mathbb{D}_a} \|(I - R^*_\gamma(z))^{-1}\| \ll 1.$$

The relation $L^*(z) = A^*(z)T^*(z)D^*(z)+E^*(z)$ from Section 4 holds in the presence of γ (with the obvious definitions) and it is immediate that

$$A^*_\gamma(z) = \gamma A^*(z), \quad D^*_\gamma(z) = D^*(z), \quad E^*_\gamma(z) = E^*(z).$$

In particular E^*_γ is a polynomial of degree at most $k-1$. By Corollary 4.5(b,c), $\sup_{z\in\mathbb{D}_a} \|A^*(z)\| \ll 1 + a^r S_r(k,a)$, and $\sup_{z\in\mathbb{D}_a} \|D^*(z)\| \ll 1 + a^r S_r(k,a)$. Hence $\sup_{z\in\mathbb{D}_a} \|L^*_\gamma(z)\| \ll (1 + a^r S_r(k,a))^2$ and the result follows. ∎

Remark B.7 The spectral radius property for $R_\gamma(z)$ holds in $L^1(Y)$, so it is possible to prove Lemma B.6 without passing to the Lipschitz norm. However, this does not seem to lead to improvements in our final results.

Proof of Theorem B.1 Suppose without loss that v is mean zero. Let $\ell \geq 1$, and write

$$\rho^*(n) = \int_{\Delta^*} v\, w \circ f^{*n}\, d\mu^* = \int_{\Delta^*} v \circ f^{*\ell}\, w \circ f^{*\ell+n}\, d\mu^* = I_1 + I_2 + I_3,$$

where

$$\begin{aligned}
I_1 &= \int_{\Delta^*} (v \circ f^{*\ell} - \tilde{v}_\ell)\, w \circ f^{*\ell+n}\, d\mu^* \\
I_2 &= \int_{\Delta^*} \tilde{v}_\ell\, (w \circ f^{*n/2} - \tilde{w}_{n/2}) \circ f^{*\ell+n/2}\, d\mu^* \\
I_3 &= \int_{\Delta^*} \tilde{v}_\ell\, \tilde{w}_{n/2} \circ f^{*\ell+n/2}\, d\mu^*.
\end{aligned}$$

By Proposition B.5(b), $|I_1| \leq |v \circ f^{*\ell} - \tilde{v}_\ell|_\infty |w|_\infty \leq C|\gamma^{\psi^*_n}\|v_0\|_{C^\eta}|w_0|_\infty \leq C\gamma^{\ell/k}\|v_0\|_{C^\eta}|w_0|_\infty$. By Proposition B.5(a) and Lemma B.6, $|I_2| \leq |\tilde{v}_\ell|_\infty |w \circ f^{*n/2} - \tilde{w}_{n/2}|_1 \ll |v_0|_\infty \|w_0\| e^{-\frac{1}{2}na(k)}$. Assume for the moment that $\tilde{v}_\ell$ is mean zero. By Theorem 4.2 and Proposition B.5(c),

$$\begin{aligned}
|I_3| &= \Big|\int_{\bar{\Delta}^*} \bar{v}_\ell\, \bar{w}_{n/2} \circ \bar{f}^{*\ell+n/2}\, d\mu^*\Big| = \Big|\int_{\bar{\Delta}^*} L^{*n/2} L^{*\ell} \bar{v}_\ell\, \bar{w}_{n/2}\, d\mu^*\Big| \\
&\leq |L^{*n/2}L^{*\ell}\bar{v}_\ell|_1 |\bar{w}_{n/2}|_\infty \ll S_q(k,a)e^{-\frac{1}{2}na}\|L^{*\ell}\bar{v}_\ell\|_\theta |w_0|_\infty \ll S_q(k,a)e^{-\frac{1}{2}na}\|v_0\|_{C^\eta}|w_0|_\infty.
\end{aligned}$$

In the general case where $\tilde{v}_\ell$ is not mean zero, we apply the above argument with $\tilde{v}_\ell$ replaced by $\tilde{v}_\ell - \int_{\Delta^*} \tilde{v}_\ell\, d\mu^*$, and there is an extra term bounded by $|\int_{\Delta^*} \tilde{v}_\ell\, d\mu^*||w_0|_\infty$. Since v is mean zero, $|\int_{\Delta^*} \tilde{v}_\ell\, d\mu^*| = |\int_{\Delta^*}(\tilde{v}_\ell - v \circ f^{*\ell})\, d\mu^*| \leq C\|v_0\|_{C^\eta} e^{-\ell/k}$ by another application of Proposition B.5(b).

Finally, ℓ is arbitrary, and letting $\ell \to \infty$ yields the result. ∎

Acknowledgements The research of IM and DT was supported in part by EPSRC Grant EP/F031807/1. We are very grateful to Sébastien Gouëzel, Stefano Luzzatto and Sandro Vaienti for helpful discussions and encouragement. Special thanks to Sébastien Gouëzel for showing IM the additional ideas [3, 7] required for Theorem B.1.

References

[1] J. Aaronson, M. Denker, O. Sarig and R. Zweimüller. Aperiodicity of cocycles and conditional local limit theorems. *Stoch. Dyn.* **4** (2004) 31–62.

[2] V. Araújo, S. Luzzatto and M. Viana. Invariant measure for interval maps with critical points and singularities. *Adv. Math.* **221** (2009) 1428–1444.

[3] J.-R. Chazottes and S. Gouëzel. Optimal concentration inequalities for dynamical systems. *Commun. Math. Phys.* **316** (2012) 843–889.

[4] W. Feller. *An Introduction to Probability Theory and its Applications, II.* Wiley, New York, 1966.

[5] S. Gouëzel. Sharp polynomial estimates for the decay of correlations. *Israel J. Math.* **139** (2004) 29–65.

[6] S. Gouëzel. Vitesse de décorrélation et théorèmes limites pour les applications non uniformément dilatantes. Ph. D. Thesis, Ecole Normale Supérieure, 2004.

[7] S. Gouëzel. Private communication.

[8] M. Holland. Slowly mixing systems and intermittency maps. *Ergodic Theory Dynam. Systems* **25** (2005) 133–159.

[9] H. Hu and S. Vaienti. Absolutely continuous invariant measures for non-uniformly expanding maps. *Ergodic Theory Dynam. Systems* **29** (2009) 1185–1215.

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

[11] S. Luzzatto and I. Melbourne. Statistical properties and decay of correlations for interval maps with critical points and singularities. *Commun. Math. Phys.* To appear.

[12] V. Maume-Deschamps. Projective metrics and mixing properties on towers. *Trans. Amer. Math. Soc.* **353** (2001) 3371–3389.

[13] I. Melbourne. Decay of correlations for slowly mixing flows. *Proc. London Math. Soc.* **98** (2009) 163–190.

[14] I. Melbourne and D. Terhesiu. Operator renewal theory and mixing rates for dynamical systems with infinite measure. *Invent. Math.* **189** (2012) 61–110.

[15] W. Parry and M. Pollicott. *Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics.* Astérique **187-188**, Société Mathématique de France, Montrouge, 1990.

[16] M. Rychlik. Bounded variation and invariant measures. *Studia Math.* **76** (1983) 69–80.

[17] O. M. Sarig. Subexponential decay of correlations. *Invent. Math.* **150** (2002) 629–653.

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

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

[20] R. Zweimüller. Ergodic structure and invariant densities of non-Markovian interval maps with indifferent fixed points. *Nonlinearity* **11** (1998) 1263–1276.