An adaptive subdivision technique for the approximation of attractors and invariant measures. Part II: Proof of convergence

Oliver Junge
Department of Mathematics and Computer Science
University of Paderborn
33098 Paderborn, Germany
junge@upb.de

March 19, 2002

Abstract

We prove convergence of a recently introduced adaptive multilevel algorithm for the efficient computation of invariant measures and attractors [Dellnitz and Junge, 1998] of dynamical systems. The proof works in the context of (sufficiently regular) stochastic processes and essentially shows that the discretization of phase space leads to a small random perturbation of the original process. A generalized version of a Lemma of Khasminskii then gives the desired result.

Key words. invariant measure, subdivision algorithm, adaptive partition, Ulam's method.

2000 AMS subject classification. 37M25 (Primary), 65P99 (Secondary)

1 General Introduction

The statistics of the long time behavior of a dynamical system are captured by certain *invariant measures*. These measures weight a part of the state space according to the probability by which a typical trajectory of the system is being observed in this part. In this sense these *natural invariant* (or SRB-) measures are objects of fundamental interest in the study of the long time behavior of dynamical systems.

In easy cases, an approximation to one of these natural invariant measures can be obtained by computing a histogram of a (very) long trajectory of the system: one counts the number of times that the trajectory enters a certain part of phase space. Unfortunately this methods lacks results on the connection between the length of the orbit and the accuracy of the approximated invariant measure.

An alternative approach, commonly referred to as "Ulam's method" is to partition (part of) the phase space into a finite number of sets and to employ the dynamical system under consideration to compute *transition probabilities* between the elements of the partition. This yields a Markov model of the original system in the form of a stochastic matrix. An approximate invariant measure is then given by an invariant vector of this matrix.

In this paper we give a proof of convergence for an adaptive version of Ulam's method. The idea of the algorithm [Dellnitz and Junge, 1998] is to use the information provided by the invariant measure to selectively refine part of the partition in one step of the algorithm. After a certain number of steps an adaptive partition has been constructed and typically the estimate of the invariant measure is significantly better than one based on a uniform partition of the same cardinality – see the example in Section 4.

The algorithm described here has been implemented and is available within the software package GAIO[1].

2 Background

In [Ulam, 1960] Ulam proposed the following method for an approximate computation of an invariant measure of a map $f : [0,1] \to [0,1]$: Divide the unit interval into k intervals $I_1, \dots, I_k$ of equal length and consider the matrix

$$(p_{ij}) = \frac{m(I_j \cap f^{-1}(I_i))}{m(I_j)},$$

$i, j = 1, \dots, k$, of *transition probabilities* between the intervals. This matrix has a largest eigenvalue 1 and Ulam conjectured that for certain maps the step function h_k defined by the eigenvector to this eigenvalue provides an approximation to the density of an invariant measure of f. In [Li, 1976] Li showed that indeed, if f is a piecewise expanding C^2-map possessing a unique invariant density h,

[1] `http://www.upb.de/math/~agdellnitz/gaio`

then h_k converges in L^1 to h as $k \to \infty$. He exploited the fact that the transition matrix arises from a discretization of the *Frobenius-Perron operator*, which for this particular class of maps can be shown to be sufficiently regular (see e.g. [G. Keller, 1982]). In the latter this result has been generalized in various directions, see e.g. [Boyarsky and Lou, 1991, Chiu et al., 1992, Boyarsky et al., 1994, Miller, 1994, Froyland, 1995, Ding and Zhou, 1996, Froyland, 1996, Hunt, 1996, Murray, 1997, Blank and Keller, 1998, Dellnitz and Junge, 1999].

In order to improve the efficiency of the method, Ding, Du and Li suggested in [Ding et al., 1993] to use ansatzfunctions of higher order instead of piecewise constant ones. However in order to get a higher rate of convergence this approach essentially requires the unknown invariant density to exhibit the same smoothness properties as the ansatzfunctions. In [Dellnitz and Junge, 1998] another property of Ulam's original method was modified. Instead of requiring the partition sets to be of equal size it was proposed to locally adapt the size of the partition sets in a suitable way to the invariant measure. To be more precise the therein advocated *adaptive subdivision algorithm* used the information provided by an approximate invariant measure in each step to selectively refine part of the partition for the next step. The numerical experiments in this paper suggested that in certain cases this algorithm delivers a higher rate of convergence than Ulam's "standard" method.

The aim of this paper is to give a convergence result for this new algorithm in the context of sufficiently regular stochastic transition functions. The advantage of considering transition functions is that as in [Dellnitz and Junge, 1999] the convergence result can directly be combined with results on the stochastic stability of (e.g.) SRB-measures of uniformly hyperbolic systems [Kifer, 1986] in order to obtain convergence results for these systems. Now a crucial property of the approach in [Dellnitz and Junge, 1999] (and the conventional "Ulam type" methods) is that the underlying partition of the phase space is uniformly refined in order to obtain convergence. Since this is the very property missing in the adaptive approach the proof in [Dellnitz and Junge, 1999] does not directly carry over.

The proof we give here (which is also contained in the thesis [Junge, 1999]) centers around a generalization of a Lemma due to Khasminskii [Khasminskii, 1963]: Consider a *small random perturbation* of the original deterministic system f given by a family of stochastic transition functions p_ε. For every ε let μ_ε be an invariant measure of p_ε. If a sequence $(\mu_{\varepsilon_k})_k$ with $\varepsilon_k \to 0$ as $k \to \infty$ converges (weakly), then the limit is an invariant measure for f. We are going to generalize the notion of a small random perturbation to the context of stochastic transition functions and prove the analogue to Khasminskii's result. Now in order to prove convergence of the adaptive subdivision algorithm we show that the discretization of the underlying stochastic process can be interpreted as a small random perturbation in the generalized sense. Then the application of (the generalized version of) Khasminskii's Lemma leads to the desired result.

The paper is organized as follows: after some introductory remarks in §3 on the discretization of the Frobenius-Perron operator in our context we describe the adaptive subdivision algorithm in §4. In order to make the paper self-contained and to give motivation for considering this particular algorithm we recall numerical results obtained in [Dellnitz and Junge, 1998] in §4. Finally in §5 we state and prove the convergence theorem.

3 Discretization of the Frobenius-Perron operator

We consider an evolution on the compact state space $X \subset \mathbb{R}^n$ defined by a transition function $p : X \times \mathcal{A} \to [0, 1]$ ($\mathcal{A}$ the Borel-σ-Algebra on X): For every $x \in X$ the function $p(x, \cdot)$ is a probability measure on X and we assume that p is *strong Feller*, i.e. that the function $x \mapsto \int g(y)\, p(x, dy)$ is continuous for every bounded, measurable function $g : X \to \mathbb{R}$. We suppose furthermore that X is forward invariant in the sense that $p(x, X) = 1$ for all x. The motivation to consider transition functions (instead of some deterministic system itself) stems from the fact that there are well established results on the stochastic stability of SRB-measures (see e.g. [Kifer, 1986]). In these theorems the stochastically perturbed deterministic system is modelled by a (sufficiently regular) stochastic transition function.

An *invariant measure* of a transition function p is a probability measure μ on X such that $\mu(A) = \int p(x, A)\, \mu(dx)$ for all measurable A. It is immediate that if we define the *Frobenius-Perron* operator $P : \mathcal{M} \to \mathcal{M}$ on the space $\mathcal{M}$ of probability measures on X by

$$P\mu(A) = \int p(x, A)\, \mu(dx), \quad A \in \mathcal{A},$$

then a measure is invariant iff it is a fixed point of P.

Now the basic idea underlying the approximate computation of invariant measures is to discretize the *fixed point problem* $P\mu = \mu$ and to look for fixed points of the discretized operator (i.e. eigenvectors of the eigenvalue 1 of a corresponding finite dimensional matrix). To this end we introduce spaces of *discrete probability measures* as follows. Let $\mathcal{B} = \{B_1, \ldots, B_b\}$ be a finite partition (of a part of X), i.e. a finite collection of compact subsets of X such that $m(B_i \cap B_j) = 0$ for $i \neq j$, where m denotes Lebesgue measure on X. The space $\mathcal{M}_\mathcal{B}$ of discrete probability measures on $\mathcal{B}$ is the set of probability measures having a density (with respect to m) which is piecewise constant on the elements of the partition $\mathcal{B}$. Obviously a discrete probability measure μ is determined by a nonnegative vector $u \in \mathbb{R}^b$ with $\|u\|_1 = 1$, where $u_i = \mu(B_i)$, $i = 1, \ldots, b$. Using the projection $Q_\mathcal{B} : \mathcal{M} \to \mathcal{M}_\mathcal{B}$,

$$Q_\mathcal{B}\mu(A) = \sum_{B \in \mathcal{B}} \frac{m(A \cap B)}{m(B)} \mu(B),$$

we define the *discretized Frobenius-Perron operator* $P_\mathcal{B} : \mathcal{M}_\mathcal{B} \to \mathcal{M}_\mathcal{B}$, $P_\mathcal{B} = Q_\mathcal{B} P$.

4 The adaptive subdivision algorithm

The algorithm described below computes a sequence of pairs $(\mathcal{B}_0, \mu_0), (\mathcal{B}_1, \mu_1), \ldots$, where each $\mathcal{B}_k$ is a partition (of a subset of X) and each μ_k is a discrete probability measure invariant under $P_{\mathcal{B}_k}$. Given $(\mathcal{B}_0, \mu_0)$, one inductively obtains $(\mathcal{B}_k, \mu_k)$ from $(\mathcal{B}_{k-1}, \mu_{k-1})$ for $k = 1, 2, \ldots$ in two steps:

(i) *Subdivision:* Choose $\delta_{k-1} \in [0, 1]$, set $\mathcal{B}^+_{k-1} = \{B \in \mathcal{B}_{k-1} : \mu_{k-1}(B) \geq \delta_{k-1}\}$ and construct a new collection $\hat{\mathcal{B}}^+_k$ by refining the elements of $\mathcal{B}^+_{k-1}$ such that $\operatorname{diam}(\hat{\mathcal{B}}^+_k) \leq \theta \operatorname{diam}(\mathcal{B}^+_{k-1})$ for some $0 < \theta < 1$.

(ii) *Selection:* Set $\hat{\mathcal{B}}_k = \hat{\mathcal{B}}^+_k \cup (\mathcal{B}_{k-1} \backslash \mathcal{B}^+_{k-1})$. Compute a fixed point μ_k of the discretized Frobenius Perron operator $P_{\hat{\mathcal{B}}_k}$. Set $\mathcal{B}_k = \{B \in \hat{\mathcal{B}}_k : \mu_k(B) > 0\}$.

We terminate the process if $\min_{B \in \mathcal{B}_k} \operatorname{diam}(B) < \beta$ for some k, where $\beta > 0$ is some prescribed minimal diameter. For details concerning the numerical realization of the algorithm we refer to [Dellnitz and Junge, 1998].

Numerical example. We recall a numerical result from [Dellnitz and Junge, 1998]. Consider the logistic map $f(x) = 4x(1-x)$ with invariant density $h_{log}(x) = (\pi\sqrt{x(1-x)})^{-1}$ (see e.g. [Lasota and Mackey, 1994]). We approximate h_{log} by Ulam's standard method using equipartitions and by applying the adaptive subdivision algorithm, choosing $\delta_k = 1/b_k$, where b_k is the number of boxes in $\mathcal{B}_k$.

The transition probabilities have been computed by the method presented in [Guder et al., 1997], for the computation of the invariant probability vector we used an Arnoldi method [Lehoucq et al., 1998] as implemented in the `eigs` command in MATLAB.

Figure 1 shows the L^1-error between h_{log} and the approximate densities versus the cardinality of the partitions. Note that since h_{log} is known the error between the approximate densities and the true one can be computed exactly.

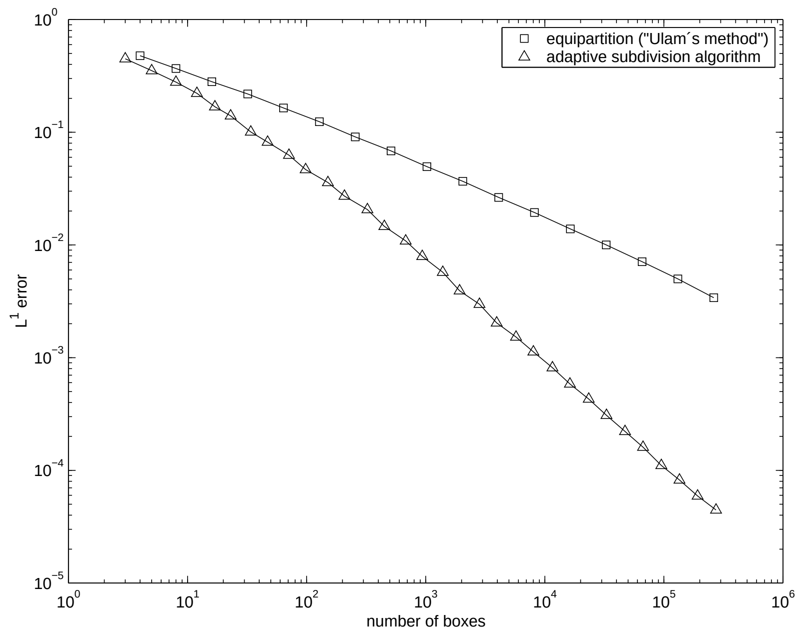

Figure 1: L^1-error between h_{log} and the approximate invariant densities versus the cardinality of the underlying partitions

5 Proof of convergence

In this section we are going to establish the following convergence result for the adaptive subdivision algorithm.

Theorem 5.1. *Let the transition function $p(x, \cdot) : \mathcal{A} \to [0, 1]$ be absolutely continuous for all x with transition density $h(x, \cdot) : X \to \mathbb{R}$ and let h be uniformly bounded in x, i.e. $h(x, y) \leq K$ for all $x, y \in X$. If p has a unique invariant measure μ and if $\delta_k \to 0$ as $k \to \infty$ then the sequence μ_k of discrete probability measures generated by the algorithm converges weakly to the invariant measure of p.*

Note in particular that we require the invariant measure of p to be unique, sufficient conditions for this can be found e.g. in [Hunt, 1996]. As stated in the introduction the main idea of the proof of theorem 5.1 is to show that the discretized Frobenius-Perron operator emerges as the transition operator of a transition function which (under the hypotheses of the theorem) can be shown to be a small random perturbation of the original transition function p. A generalized version of a Lemma of Khasminskii then immediately gives the desired result.

So let us begin with the generalized notion of a small random perturbation. For ease of notation we are going to write

$$\mu(g) = \int g \, d\mu \quad \text{and} \quad p(x,g) = \int g(y)\, p(x,dy)$$

for measures μ and measurable, bounded functions $g : X \to \mathbb{R}$.

Definition 5.2. *Let p be a transition function and p_κ, $\kappa > 0$, be a family of transition functions. If there is a family of probability measures μ_κ, such that*

$$\int |p(x,g) - p_\kappa(x,g)| \, \mu_\kappa(dx) \to 0$$

as $\kappa \to 0$ for continuous functions $g : X \to \mathbb{R}$, then the family p_κ is called a small random perturbation *of p (with respect to the family μ_κ).*

It is immediate to generalize the Lemma of Khasminskii [Khasminskii, 1963] to this context.

Lemma 5.3. *Consider a small random perturbation p_κ of a transition function p with respect to a family μ_κ of probability measures. Suppose that p is strong Feller. If μ_κ is an invariant measure of p_κ for each κ and $\mu_{\kappa_i} \to \mu$ weakly for some subsequence $\kappa_i \to 0$ and $i \to \infty$, then μ is an invariant measure for p.*

Proof. Let $g : X \to \mathbb{R}$ be a continuous function. Then

$$\begin{aligned} |\mu(g) - (P\mu)(g)| &\leq |\mu(g) - \mu_{\kappa_i}(g)| + |(P_{\kappa_i}\mu_{\kappa_i})(g) - (P\mu_{\kappa_i})(g)| \\ &+ |(P\mu_{\kappa_i})(g) - (P\mu)(g)|. \end{aligned}$$

The first term on the right hand side vanishes for $i \to \infty$ since we supposed μ_{κ_i} to converge weakly to μ. Let us consider the third term. We have $(P\mu)(g) = \int p(x,g)\ \mu(dx)$. Since p was supposed to be strong Feller, the map $x \mapsto p(x,g)$ is continuous. Using the weak convergence of μ_{κ_i} again, it follows that the third term vanishes as $i \to \infty$ as well. Finally we use the fact that p_κ is a small random perturbation of p in order to see that also the second term goes to zero.

Thus we get $\mu(g) = (P\mu)(g)$, and this relation being true for any continuous function g implies that μ is an invariant measure for p. □

The next step in the proof of theorem 5.1 is to relate the discretized Frobenius-Perron operator to some transition function which in turn will then be shown to be a small random perturbation of p. In order to find a suitable transition function we have a closer look at the discretized Frobenius-Perron operator:

$$(P_\mathcal{B}\mu)(A) \;=\; (Q_\mathcal{B}P\mu)(A) = \int \sum_{B\in\mathcal{B}} \frac{m(A\cap B)}{m(B)} p(x,B)\;\mu(dx) \tag{5.1}$$

Thus $P_\mathcal{B}$ is the transition operator of the transition function

$$p_\mathcal{B}(x,A) = \sum_{B\in\mathcal{B}} \frac{m(A\cap B)}{m(B)} p(x,B) = \int \mathbb{1}_A(y) \sum_{B\in\mathcal{B}} \frac{\mathbb{1}_B(y)}{m(B)} p(x,B)\; m(dy)$$

which will be the object of main interest further on.

Our goal now is to show that (under the hypotheses of theorem 5.1) the family $p_{\mathcal{B}_k}$ is a small random perturbation of p with respect to the family μ_k (where $\mathcal{B}_k, \mu_k$ are the partitions and measures generated by the adaptive subdivision algorithm). According to definition 5.2 we have to show that, roughly speaking, asymptotically the discrete transition distribution $p_{\mathcal{B}_k}(x,\cdot)$ equals $p(x,\cdot)$ *for ν-almost all x*, where ν is an accumulation point of the sequence μ_k. It is crucial to note that we do not have to show asymptotic equality of $p_{\mathcal{B}_k}(x,\cdot)$ and $p(x,\cdot)$ for all x, not even for Lebesgue almost all x. Since at the end the accumulation point ν turns out to be an invariant measure of p, we are in some sense using the information provided by the invariant measure ν beforehand.

So we have to show that

$$\int |p(x,g) - p_{\mathcal{B}_k}(x,g)|\; \mu_k(dx) \to 0 \tag{5.2}$$

for $k \to \infty$ and continuous functions g. Note that we can write

$$\int g(y)\; p_{\mathcal{B}_k}(x,dy) = \int (Q_{\mathcal{B}_k} g)(y)\; p(x,dy),$$

where we define the projection $Q_{\mathcal{B}_k}$ from L^1 into the space of simple functions on the partition $\mathcal{B}_k$ by $Q_{\mathcal{B}_k} g = \sum_{B\in\mathcal{B}_k} \frac{\mathbb{1}_B}{m(B)} \int_B g\; dm$. Using this rewrite we estimate

$$\begin{aligned} |p(x,g) - p_{\mathcal{B}_k}(x,g)| \;&\le\; \int |g(y) - (Q_{\mathcal{B}_k} g)(y)|\; p(x,dy) \\ &=\; p(x, |g - Q_{\mathcal{B}_k} g|) \end{aligned} \tag{5.3}$$

So abbreviating $d_k = |g - Q_{\mathcal{B}_k} g|$ what we finally have to show is that

$$\int p(x, d_k)\, \mu_k(dx) \to 0 \tag{5.4}$$

for $k \to \infty$ and every continuous function g.

To this end we split up the partitions $\mathcal{B}_k$ into a "transient" part $\mathcal{T}_k$ and a "supporting" part $\mathcal{S}_k$. In $\mathcal{T}_k$ we collect all partition sets which never get subdivided again or have been removed in a previous selection step of the algorithm, i.e.

$$\mathcal{T}_k = \{B \in \mathcal{B}_k : \mu_\ell(B) < \delta_\ell \text{ for } \ell \geq k\} \,\cup\, \bigcup_{j=1}^{k} \hat{\mathcal{B}}_j \backslash \mathcal{B}_j.$$

Accordingly let $\mathcal{S}_k = \mathcal{B}_k \backslash \mathcal{T}_k$. We denote $T_k = \bigcup_{B \in \mathcal{T}_k} B$, $\quad T = \bigcup_{k=0}^{\infty} T_k$, $\quad S_k = \bigcup_{B \in \mathcal{S}_k} B$ and $S = X \backslash T$. We are now going to show (5.4) by splitting up the integral into two parts: $p(x, d_k) = p(x, \mathbb{1}_{T_k} d_k) + p(x, \mathbb{1}_{S_k} d_k)$.

The transient part T. Let ν be an accumulation point of the sequence $(\mu_k)_k$. In the following we treat a corresponding convergent subsequence which, abusing notation, we will also denote by $(\mu_k)_k$. We treat the transient part by showing that the transition probability into T is zero for ν-almost all points.

Proposition 5.4. *If $\delta_\ell \to 0$ as $\ell \to \infty$ then $p(x, T) = 0$ for ν-almost all $x \in X$.*

Proof. It suffices to show the assertion for any $B \in \mathcal{T}_k$, $k = 0, 1, \ldots$, instead of T. Since μ_ℓ is a fixed point of $P_{\mathcal{B}_\ell}$ for all $\ell = 0, 1, \ldots$, we have (using equation (5.1))

$$\mu_\ell(B) \;=\; (P_{\mathcal{B}_\ell} \mu_\ell)(B) = \int p(x, B)\, \mu_\ell(dx)$$

Since p is strong Feller the function $x \mapsto p(x, B)$ is continuous. Thus

$$\mu_\ell(B) \to \int p(x, B)\, \nu(dx) \quad \text{as } \ell \to \infty,$$

taking a suitable subsequence if necessary. Now using the fact that $B \in \mathcal{T}_k$, i.e. $\mu_\ell(B) < \delta_\ell$ for $\ell \geq k$, and the assumption $\delta_\ell \to 0$ ($\ell \to \infty$) we obtain the assertion of the proposition. $\square$

Observe that for every continuous function g there is constant $C > 0$ such that $d_k \leq C$ for all k. Thus for all $\varepsilon > 0$ there is a $K = K(\varepsilon) \in \mathbb{N}$ such that for all $k \geq K$

$$\begin{aligned} \int p(x, \mathbb{1}_T d_k)\, \mu_k(dx) \;&\leq\; C \int p(x, T)\, \mu_k(dx) \\ &\leq\; C \left(\int p(x, T)\, \nu(dx) + \varepsilon \right) = C\varepsilon. \end{aligned}$$

Combining this with the fact that $p(x, \mathbb{1}_{T_k} d_k) \leq p(x, \mathbb{1}_T d_k)$ we conclude that

$$\int p(x, \mathbb{1}_{T_k} d_k)\ \mu_k(dx) \to 0$$

as $k \to \infty$, finishing the proof for the transient part.

The supporting part S. In order to prove convergence on S we are going to use the fact that the diameter of the collections $\mathcal{S}_k$ is going to zero as k tends to infinity. Furthermore, since $S_{k+1} \subset S_k$, S can be viewed as the limit of the sets S_k, i.e. $S = \bigcap_{k=0}^{\infty} S_k$.

In the situation that we are interested in the transition function p is absolutely continuous. In this case it is easy to see that S has positive Lebesgue-measure. Suppose this was not the case, i.e. $m(S) = 0$. Then we had $p(x, S) = 0$ for all x by the absolute continuity of p and $p(x, T) = 0$ for ν-almost all x according to Proposition 5.4. Since $X = S \cup T$ this is a contradiction to the fact that $p(x, \cdot)$ is a probability measure.

Proposition 5.5.

$$\int_{S_k} d_k\ dm \to 0 \quad \textit{for } k \to \infty.$$

Proof. If we define for a bounded measurable function $g : X \to \mathbb{R}$

$$Q_k g = \sum_{B \in \mathcal{S}_k} \frac{\mathbb{1}_B}{m(B)} \int_B g\ dm,$$

we get

$$\begin{aligned} \int_{S_k} d_k\ dm &= \int |Q_{\mathcal{B}_k} g - g| \mathbb{1}_{S_k}\ dm = \int |Q_k g - g\mathbb{1}_{S_k}|\ dm \\ &\leq \int |Q_k g - g\mathbb{1}_S|\ dm + \int |g\mathbb{1}_S - g\mathbb{1}_{S_k}|\ dm \end{aligned} \tag{5.5}$$

For the second integral in (5.5) we get

$$\int |g\mathbb{1}_S - g\mathbb{1}_{S_k}|\ dm \leq \max_{x \in X} |g(x)|\ m(S_k \backslash S) \to 0$$

for $k \to \infty$. The first integral in (5.5) can be bounded as follows:

$$\begin{aligned} \int |Q_k g - g\mathbb{1}_S|\ dm &\leq \int |Q_k g - Q_k(g\mathbb{1}_S)|\ dm \\ &+ \int |Q_k(g\mathbb{1}_S) - g\mathbb{1}_S|\ dm. \end{aligned} \tag{5.6}$$

For the first integral on the right hand side of (5.6) we get

$$\begin{aligned}\int |Q_k g - Q_k(g\mathbb{1}_S)| \; dm &= \int \left| \sum_{\substack{B\in\mathcal{S}_k\\ B\cap T\neq\emptyset}} \frac{\mathbb{1}_B}{m(B)} \int_B g - g\mathbb{1}_S \; dm \right| \; dm \\ &\leq \int \sum_{\substack{B\in\mathcal{S}_k\\ B\cap T\neq\emptyset}} \frac{\mathbb{1}_B}{m(B)} \left(\int_{B\setminus S} |g| \; dm \right) dm \\ &= \sum_{\substack{B\in\mathcal{S}_k\\ B\cap T\neq\emptyset}} \int_{B\setminus S} |g| \; dm \leq \max_{x\in X} |g(x)| \; m(S_k\setminus S) \to 0\end{aligned}$$

for $k \to \infty$. For the second integral on the right hand side of (5.6) we use standard arguments from measure theory (see e.g. [Mañé, 1987]). We enlarge the collection $\mathcal{S}_k$ to a partition $\hat{\mathcal{S}}_k$ of X, requiring that $\operatorname{diam}\left(\hat{\mathcal{S}}_k\setminus\mathcal{S}_k\right) \leq \operatorname{diam}(\mathcal{S}_k)$. For $\hat{g} \in L^1$ we consider the projection

$$\hat{Q}_k \hat{g} = \sum_{B\in\hat{\mathcal{S}}_k} \frac{\mathbb{1}_B}{m(B)} \int_B \hat{g} \; dm,$$

and observe that $\hat{Q}_k \hat{g} = Q_k \hat{g}$, if $\operatorname{supp}(\hat{g}) \subset S_k$. We can now rewrite the second integral on the right hand side of (5.6) as

$$\int |Q_k(g\mathbb{1}_S) - g\mathbb{1}_S| \; dm = \int |\hat{Q}_k(g\mathbb{1}_S) - g\mathbb{1}_S| \; dm.$$

Since $\operatorname{diam}(\hat{\mathcal{S}}_k) \to 0$ we get $\|\hat{Q}_k\hat{g} - \hat{g}\|_1 \to 0$ for $k \to \infty$, this ends the proof of Proposition 5.5. □

Finally we use the last assumption of Theorem 5.1, namely that the transition density h is bounded. We get

$$\begin{aligned}\int p(x, \mathbb{1}_{S_k} d_k) \; d\mu_k &= \int \left(\int_{S_k} d_k(y) \; h(x,y) \; m(dy) \right) \; \mu_k(dx) \\ &\leq K \int_{S_k} d_k \; dm \to 0\end{aligned}$$

for $k \to \infty$, according to Proposition 5.5.

To recapitulate, we showed that $\int p(x, \mathbb{1}_{T_k} d_k) \; d\mu_k \to 0$ and $\int p(x, \mathbb{1}_{S_k} d_k) \; d\mu_k \to 0$ as $k \to \infty$, i.e.

$$\int p(x, d_k) \; d\mu_k = \int p(x, |g - Q_{\mathcal{B}_k} g|) \; d\mu_k \to 0,$$

which according to the estimate (5.3) and (5.2) shows that the transition function $p_{\mathcal{B}_k}$ is indeed a small random perturbation of p with respect to the family $(\mu_k)_k$. Using Lemma 5.3 this ends the proof of Theorem 5.1.

6 Acknowledgements

The author is grateful to Michael Dellnitz for helpful comments on the proof and to Gary Froyland for thorough proofreading of the manuscript. He would also like to thank the referees for valuable feedback.

References

[Blank and Keller, 1998] Blank, M. and Keller, G. (1998). Random perturbations of chaotic dynamical systems: stability of the spectrum. *Nonlinearity*, 11(5):1351–1364.

[Boyarsky et al., 1994] Boyarsky, A., Gora, P., and Lou, Y. (1994). Constructive approximations to the invariant densities of higher dimensional transformations. *Cons. Appr.*, 10(1):1–13.

[Boyarsky and Lou, 1991] Boyarsky, A. and Lou, Y. (1991). Approximating measures invariant under higher-dimensional chaotic transformations. *J. Appr. Theory*, 65(2):231–244.

[Chiu et al., 1992] Chiu, C., Du, Q., and Li, T. (1992). Error estimates of the Markov finite approximation of the Frobenius-Perron operator. *Nonlinear Analysis*, 19(4):291–308.

[Dellnitz and Junge, 1998] Dellnitz, M. and Junge, O. (1998). An adaptive subdivision technique for the approximation of attractors and invariant measures. *Comput. Visual. Sci.*, 1:63–68.

[Dellnitz and Junge, 1999] Dellnitz, M. and Junge, O. (1999). On the approximation of complicated dynamical behavior. *SIAM J. Numer. Anal.*, 36(2):491–515.

[Ding et al., 1993] Ding, J., Du, Q., and Li, T. Y. (1993). High order approximation of the Frobenius-Perron operator. *Appl. Math. Comp.*, 53:151–171.

[Ding and Zhou, 1996] Ding, J. and Zhou, A. (1996). Finite approximations of Frobenius-Perron operators. A solution of Ulam's conjecture to multi-dimensional transformations. *Physica D*, 1-2:61–68.

[Froyland, 1995] Froyland, G. (1995). Finite approximation of Sinai-Bowen-Ruelle measures for Anosov systems in two dimensions. *Random Comp. Dyn.*, 3(4):251–263.

[Froyland, 1996] Froyland, G. (1996). *Estimating Physical Invariant Measures and Space Averages of Dynamical Systems Indicators.* PhD thesis, University of Western Australia. Available from `http://www.upb.de/math/~froyland`.

[G. Keller, 1982] G. Keller (1982). Stochastic stability in some chaotic dynamical systems. *Monatsh. Math.*, 94:313–333.

[Guder et al., 1997] Guder, R., Dellnitz, M., and Kreuzer, E. (1997). An adaptive method for the approximation of the generalized cell mapping. *Chaos, Solitons and Fractals*, 8(4):525–534.

[Hunt, 1996] Hunt, F. (1996). Approximating the invariant measures of randomly perturbed dissipative maps. *J. Math. Anal. Appl.*, 198(2):534–551.

[Junge, 1999] Junge, O. (1999). *Mengenorientierte Methoden zur numerischen Analyse dynamischer Systeme*. PhD thesis, University of Paderborn. Available from `http://www.upb.de/math/~junge`.

[Khasminskii, 1963] Khasminskii, R. (1963). The averaging principle for parabolic and elliptic differential equations and Markov processes with small diffusion. *Theor. Prob. Appl.*, 6:1–21.

[Kifer, 1986] Kifer, Y. (1986). General random perturbations of hyperbolic and expanding transformations. *J. Analyse Math.*, 47:111–150.

[Lasota and Mackey, 1994] Lasota, A. and Mackey, M. (1994). *Chaos, Fractals and Noise*. Springer.

[Lehoucq et al., 1998] Lehoucq, R. B., Sorensen, D. C., and Yang, C. (1998). *ARPACK users' guide*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA. Solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods.

[Li, 1976] Li, T.-Y. (1976). Finite approximation for the Frobenius-Perron operator. A solution to Ulam's conjecture. *J. Approx. Theory*, 17:177–186.

[Mañé, 1987] Mañé, R. (1987). *Ergodic Theory and Differentiable Dynamics*. Springer.

[Miller, 1994] Miller, W. (1994). Stability and approximation of invariant measures for a class of nonexpanding transformations. *Nonlinear Analysis*, 23(8):1013–1025.

[Murray, 1997] Murray, R. (1997). *Discrete approximation of invariant densities*. PhD thesis, University of Cambridge.

[Ulam, 1960] Ulam, S. M. (1960). *A collection of mathematical problems.* Interscience.