On the computation of invariant measures in random dynamical systems*

Peter Imkeller
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
Germany

Peter Kloeden
Fachbereich Mathematik
J-W von Goethe-Universität
Postfach 11 19 32
60054 Frankfurt am Main
Germany

May 26, 2003

Abstract

Invariant measures of dynamical systems generated e. g. by difference equations can be computed by discretizing the originally continuum state space, and replacing the action of the generator by the transition mechanism of a Markov chain. In fact they are approximated by stationary vectors of these Markov chains. Here we extend this well known approximation result and the underlying algorithm to the setting of random dynamical systems, i.e. dynamical systems on the skew product of a probability space carrying the underlying stationary stochasticity and the state space, a particular non-autonomous framework. The systems are generated by difference equations driven by stationary random processes modelled on a metric dynamical system. The approximation algorithm involves spatial discretizations and the definition of appropriate random Markov chains with stationary vectors converging to the random invariant measure of the system.

2000 AMS subject classifications: 60 G, 60 H 25, 60 J 10, 65 C 30, 65 Q 05, 37 H 10, 37 M 25.

Key words and phrases: dynamical system; random dynamical system; approximation; invariant measure; Markov chain; random Markov chain; interval matrix; spatial discretization.

*This work was partially supported by the DFG research center 'Mathematics for key technologies' (FZT 86) in Berlin.

1 Introduction

Computational methods are a basic tool in the investigation of dynamical systems, both to explore what may happen and to approximate specific dynamical features such as limit cycles or attractors and, more generally, invariant measures; see, e.g., [5, 20, 23].

Given the intrinsic discreteness of the arithmetic field in a computer, such calculations inevitably lead to the spatial discretization of the dynamical system, although this is not always taken into account. However, even simple dynamical structures such as equilibria may not be replicated under a spatial discretization or spurious dynamics may be introduced; see, e.g., the example of the Hénon attractor under spatial discretization in [10]. These specific considerations as well as many others indicate that invariant measures are often the more natural and technically robust dynamical feature to approximate.

Invariant measures are a central concept in the theory of dynamical systems, both deterministic and random, and their investigation has been closely intertwined with developments in ergodic theory, see, e.g., [9, 18]. Just how one can approximate or compute an invariant measure of deterministic systems has mathematical roots in ergodic theory oriented papers from at least the 1970s, e.g., [22], or [15] dealing with the approximation of shift invariant measures by invariant probabilities written on finite alphabets in the context of Krieger's [19] finite generator theorem (see also Denker [7]). A variety of methods have since been proposed and implemented for computing invariant measures, see, e.g., [5, 6, 11, 17] and the references cited therein.

Thinking in terms of spatial discretization, though this may not be *per se* the aim of a particular investigation, one approximates a mapping on a continuum with a transition matrix on the discretized state space. Since the original generator may not map the discretized state space onto itself, one may have to choose randomly between several potential nearest grid points. In this way one approximates the original deterministic dynamical system given essentially by a difference equation on the continuum state space by a Markov chain moving between finitely many states obtained from a spatial discretization, even though the original system is purely deterministic. This idea was used by Diamond *et al* [12] to construct sequences of stationary probability vectors of such Markov chains to approximate an invariant measure of a deterministic discrete time dynamical system. In fact, they showed that the measures which could be approximated in this way were precisely the semi-invariant measures of the dynamical system, these coinciding with the invariant measures if the mapping generating the system is continuous.

Our aim here is to establish the analogous result for discrete time *random dynamical systems* [1], generated by a difference equation on a compact state space X which is driven by a random process modelled as a metric dynamical system, i.e. a measurable mapping θ defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. θ is $\mathbf{P}$-invariant. We further assume it to be invertible and continuous in both time directions. Although the resulting skew product system can be considered as a measurable autonomous semi-dynamical system, the results in [12] cannot be applied directly since the driving system compo-

nent remains unchanged and is not subjected to spatial discretization. However, the general thrust of the proof in [12] does carry over in the process of spatial discretization, i.e., introduction of approximating Markov chains on the state space X of the random mapping. Now our objects of interest are random measures, random Markov chains and random probability vectors, all evolving under the influence of the driving metric dynamical system. In the end our results seem just to be "ω"-wise counterparts of those in the deterministic case, and essentially we extend the deterministic Markov chain approximation *fibre by fibre* to a trivial *fibred system* in the sense of [9]. To justify this, careful measure theoretic arguments are needed throughout. For this reason, we follow closely the structure of the proof in [12].

Further extensions of our results under weaker assumptions than the ones we employ are easily conceivable. Firstly, one can think of relaxing the structure hypothesis of our system. Instead of considering skew products living on the product of an underlying probability space Ω and a fixed state space X, one could pass to *fibred systems* (see [9]) in which instead of being identical to a fixed X the fibres may depend on $\omega \in \Omega$. This is the setting of the random extension of the thermodynamic formalism of Ruelle [21], in the framework of which Markovian transfer operators may be analyzed using methods of ergodic and information theory. See the thesis of Bogenschütz [2] and the subsequent papers [3], [4]. These transfer operators also play a crucial role if another of the properties we use here is relaxed: the *invertibility of the transformation* θ. In case of finite-to-one expanding maps θ for example Ruelle's theorem for transfer operators, via the Perron-Frobenius theory, provides positive eigenfunctions and dually, probability eigenmeasures with common positive eigenvalues. In this and generalized contexts, variational principles for pressure functionals yield *relative equilibrium measures* which may replace the invariant Markov measures in the fibres by which in our framework we are able to approximate, and which in a more generalized setting might not exist. For concepts and results along these lines see [8].

Here is an outline of the organization of the paper. We describe and formulate our problem in the next section. In section 3 we introduce and obtain results for "intervals" of random stochastic matrices, which we need in the proof of our main Theorem. Section 4 presents a number of key properties of the weak topology of random measures. They are helpful in the proof of our main result (Theorem 5.3), which we state and prove in section 5.

2 Formulation of the problem

The invariant measures we shall compute are associated with random dynamical systems. These objects are thought of being generated by *random difference equations (RDE)* of the form

$$x_{n+1}(\omega) = f(\theta^n \omega, x_n(\omega)), \quad n \in \mathbf{Z}_+, \omega \in \Omega, \tag{1}$$

where $(\Omega, \mathcal{F}, \mathbf{P})$ is a probability space, with respect to which $\theta : \Omega \to \Omega$ is a bi-measurable $\mathbf{P}-$invariant mapping with inverse θ^{-1}, and

$$f : \Omega \times X \to X$$

is a measurable mapping on the compact metric space (X, d). In terms of Arnold [1] we are therefore facing a discrete time *random dynamical system* (θ, ϕ) defined over the *metric dynamical system* $(\Omega, \mathcal{F}, \mathbf{P}, \theta)$ if we define recursively

$$\begin{aligned} \phi(\omega, 0)x &= x, \\ \phi(\omega, n+1)x &= f(\theta^n\omega, \phi(\theta^n\omega, x)), \end{aligned}$$

for $x \in X, \omega \in \Omega, n \in \mathbf{Z}_+$. $\pi = (\theta, \phi)$ is also a skew product flow over $\Omega \times X$. The random mapping f induces an operator f^* on the set of random probability measures on $\mathcal{F} \otimes \mathcal{B}(X)$, where $\mathcal{B}(X)$ is the σ-algebra of Borel subsets of X. For such a probability measure μ it is defined by

$$f^*\mu(C) = \mu\left(F^{-1}(C)\right), \quad C \in \mathcal{F} \otimes \mathcal{B}(X) \tag{2}$$

where $F : \Omega \times X \to \Omega \times X$ is defined by $F(\omega, x) = (\theta\omega, f(\omega, x))$. The disintegrations of μ are given by the family of random probabilities μ_ω on $\mathcal{B}(X)$ satisfying

$$\mu(C) = \int_\Omega \mu_\omega(C_\omega) d\mathbf{P}(\omega), \quad C \in \mathcal{F} \otimes \mathcal{B}(X).$$

In terms of the disintegrations $\mu_\omega, f^*\mu_\omega, \omega \in \Omega$, of the two measures the property (2) is expressed by

$$(f^*\mu)_{\theta\omega}(B) = \mu_\omega(f^{-1}(\omega, B)), \quad \text{for} \quad B \in \mathcal{B}(X)$$

for $\mathbf{P}-$a.e. $\omega \in \Omega$, or, more stringently, due to the invariance of θ by

$$(f^*\mu)_{\theta^{n+1}\omega}(B) = \mu_{\theta^n\omega}(f^{-1}(\theta^n\omega, B)), \quad \text{for all} \quad B \in \mathcal{B}(X), n \in \mathbf{N},$$

for $\mathbf{P}-$a.e. $\omega \in \Omega$. A random probability measure μ on $\mathcal{F} \otimes \mathcal{B}(X)$ is called *invariant measure* for the random dynamical system π if $f^*\mu = \mu$, in other terms if

$$\mu(\pi^{-1}(C)) = \mu(C), \quad C \in \mathcal{F} \otimes \mathcal{B}(X).$$

We consider an invariant measure μ of π. We first remark that the invariance property has a popular counterpart in terms of the disintegrations $\mu_\omega, \omega \in \Omega$, of μ. They have to satisfy the invariance property

$$\mu_\omega(\phi^{-1}(\omega, 1)(B)) = \mu_\omega(f(\omega, \cdot)^{-1}(B)) = \mu_{\theta\omega}(B), \quad B \in \mathcal{B}(X),$$

for $\mathbf{P}-$a.e. $\omega \in \Omega$. Note that this is equivalent to the property

$$\mu_\omega(f(\theta^n\omega, \cdot)^{-1}(B)) = \mu_{\theta^{n+1}\omega}(B), \quad B \in \mathcal{B}(X), n \in \mathbf{N},$$

for $\mathbf{P}-$a.e. $\omega \in \Omega$.

To compute a fixed invariant measure μ of π, let us consider a finite discretization sequence $(X_M)_{M \in \mathbf{N}}$ of X, given by

$$X_M = \{x_1^{(M)}, \cdots, x_M^{(M)}\} \subset X,$$

with maximal step size

$$h_M = \sup_{x \in X} d(x, X_M),$$

with

$$d(x, A) = \inf_{y \in A} d(x, y)$$

for $A \subset X, M \in \mathbf{N}$. We shall actually approximate μ by means of invariant stochastic vectors associated with *random Markov chains* based on transitions between the states given by the discretizations X_M. Hence we are dealing with random $M \times M$ matrices, i. e. measurable mappings

$$p_M : \Omega \to \mathcal{S}_M,$$

where $\mathcal{S}_M$ denotes the set of $M \times M$ nonrandom stochastic matrices satisfying the property

$$p_M^n(\theta^m \omega) p_M^m(\omega) = p_M^{m+n}(\omega), \quad \omega \in \Omega, \quad m, n \in \mathbf{Z}_+, \tag{3}$$

where, for convenience, matrices in power 0 correspond to the unit matrix. Let $\Sigma_M \subset \mathbf{R}^M$ the subset of all $M-$dimensional stochastic vectors, i.e. non-negative vectors whose components sum up to 1. The discrete counterpart of our invariant measure is given by a measurable mapping

$$\overline{p} : \Omega \to \Sigma_M$$

which will be called p_M-*invariant* provided it satisfies the equation

$$p_M(\omega)\overline{p}(\omega) = \overline{p}(\theta\omega), \quad \text{for} \quad \mathbf{P} - a.e.\ \omega \in \Omega.$$

Then the basic problem we face in this paper can be formulated as follows.

Suppose the random cocycle generated by f possesses an invariant measure μ. Does there exist a sequence of random Markov chains $(p_M)_{M \in \mathbf{N}}$ converging to f as $N \to \infty$ and $h_M \to 0$ such that p_M possesses a random invariant probability vector $\overline{p}_M$ converging to μ as $N \to \infty$?

Of course, to tackle this problem, we first have to say what we mean by the two types of convergences appearing in the statement. Defining for *fixed* $\omega \in \Omega$ the distance between the Markov chain $p_M(\omega)$ and the mapping $f(\omega, \cdot)$ we recall that we have to imagine $p_M(\omega)$ as a *randomization* of $f(\omega, \cdot)$ the Markovian character of which relates to the flow property of the dynamical system generated by f. To measure the distance between random probability measures on $\mathcal{F} \otimes \mathcal{B}(X)$ we use the *Prokhorov distance* ρ defined for probability measures on our compact space (X, d) (usually defined, more generally, on a Polish space)

$$\rho(\mu_1, \mu_2) = \inf\{\epsilon > 0 : \mu_1(G_\epsilon(B)) \leq \mu_2(B) - \epsilon \quad \text{for all} \quad B \in \mathcal{B}(X)\}$$

(see Ethier and Kurtz [13]), where $G_\epsilon(S)$ stands for the $\epsilon-$neighborhood of some set $S \subset X$, i.e.

$$G_\epsilon(S) = \cup_{x \in S} G_\epsilon(x),$$

and $G_\epsilon(x)$ is the ball of radius ϵ centered at $x \in X$. So for random probability measures μ, ν on $\mathcal{F} \otimes \mathcal{B}(X)$ with $\omega-$disintegrations $\mu_\cdot, \nu_\cdot$ we define

$$R(\mu, \nu) = \mathbf{E}\left(\rho(\mu_\cdot, \nu_\cdot)\right) = \int_\Omega \rho\left(\mu_\omega, \nu_\omega\right) \mathbf{P}(d\omega) \tag{4}$$

In the associated topology, due to the invariance of θ, a sequence of random probability measures $(\mu^{(k)})_{k\in\mathbf{N}}$ converges to μ iff the sequence of random variables $(\rho(\mu^{(k)}_{\theta^n \cdot}, \mu_{\theta^n \cdot}))_{k\in\mathbf{N}}$ converges to 0 in $L^1(\Omega, \mathcal{F}, \mathbf{P})$ for all $n \in \mathbf{Z}$. In a similar way we define the distance of a random Markov chain $p : \Omega \to \mathcal{S}_M$ and the generator f of our random dynamical system. We let

$$D(p, f) = \sum_{i,j=1}^{M} \mathbf{E}\left(p_{ij}(\cdot)\mathrm{dist}_{X\times X}((x_i^{(M)}, x_j^{(M)}), \mathrm{Gr} f(\cdot, \cdot))\right), \tag{5}$$

where the distance on the random graph is given by

$$\mathrm{dist}_{X\times X}((x, y), \mathrm{Gr} f) = \inf_{z\in X} \max\{d(x, z), d(y, f(\cdot, z))\}, \quad x, y \in X,$$

and also write $D(p(\omega), f(\omega, \cdot))$ to measure the distance between matrices and mappings for fixed $\omega \in \Omega$, by erasing the expectation from the above formula. With respect to this notion, a sequence of Markov chains $(p^{(k)})_{k\in\mathbf{N}}$ converges to f iff for all $n \in \mathbf{Z}$ we have $(p^{(k)}(\theta^n \cdot))_{k\in\mathbf{N}}$ converges to $f(\theta^n \cdot, \cdot)$ in $L^1(\Omega, \mathcal{F}, \mathbf{P})$, and the $\omega-$wise convergence is in the sense of [12]. In this case we call the sequence of random Markov chains *approximating sequence* for the random dynamical system ϕ. We shall often have to state equations or inequalities between random variables or vectors. If not made precise, this is always tacitly assumed to mean $\mathbf{P}-$a.s. equality or inequality.

Remark: Instead of random difference equations of the form (1) one is often interested in situations where domain and image spaces of the mapping $f(\omega, \cdot)$ also depend on ω, that is, with fibred spaces X_ω such that $f(\omega, \cdot) : X_\omega \to X_{\theta\omega}$ for each $\omega \in \Omega$, see, e.g., [2, 9, 16]. In our context the measure μ would now be defined on $\cup_{\omega\in\Omega}\left(\{\omega\} \times X_\omega\right)$ and the disintegrated measures μ_ω on X_ω. In addition, we would need grids $X_M(\omega)$ $= \{x_1^{(M)}(\omega), \ldots, x_1^{(M)}(\omega)\}$ of the same finite cardinality M and with a common deterministic upper bound h_M on the maximal stepsizes; typically the fibres X_ω will be subspaces of a common metric space X and we could use a common determinsitic grid $X_M = \{x_1^{(M)}, \ldots, x_1^{(M)}\}$ on X with transition probability $p_{ij}(\omega)$ set to zero if either $x_i^{(M)} \notin X_\omega$ or $x_j^{(M)} \notin X_{\theta\omega}$. However, we will restrict our proofs here to the case formulated above with a common state space X in order not to further complicate an already technically complicated situation. For the same reasons, we will not consider the extension to case of a noninvertible θ here, referring the reader to [9] for hints as to how this could be done.

3 Intervals of random stochastic matrices

Before discussing the question of approximation of invariant measures just sketched we have to treat some preliminaries on the ordering of random stochastic matrices.

We shall introduce an ordering on the set $\mathcal{M}$ of real $M \times M$ matrices $A = (a_{ij}) = (a_{ij})_{i \le i,j \le M}$ with non-negative entries. According to the natural partial ordering on $\mathcal{M}$, we have

$$A \le B \quad \text{iff} \quad a_{ij} \le b_{ij} \quad \text{for} \quad 1 \le i,j \le M,$$

for $A, B \in \mathcal{M}$. Denote by $\mathcal{M}^-$ the set of all matrices in $\mathcal{M}$ for which the column sums do not exceed 1, i.e. $A = (a_{ij})$ such that

$$\sum_{i=1}^{M} a_{ij} \le 1, \quad 1 \le j \le M,$$

and by $\mathcal{M}^+$ accordingly those characterized by the condition

$$\sum_{i=1}^{M} a_{ij} \ge 1, \quad 1 \le j \le M.$$

Then the stochastic $M \times M$ matrices $\mathcal{S}_M$ are given by $\mathcal{M}^- \cap \mathcal{M}^+$. For $A \in \mathcal{M}^-, B \in \mathcal{M}^+$ we define

$$\widehat{AB} = \{C \in \mathcal{S}_M : A \le C \le B\}.$$

Recall the simplex of stochastic $M-$vectors is denoted by Σ_M.

The matrices we deal with are random, i.e. measurable maps $A : \Omega \to \mathcal{M}$. For random matrices A, B we write

$$A \le B \quad \text{iff} \quad A(\omega) \le B(\omega) \quad \text{for} \quad \mathbf{P} - a.e. \quad \omega \in \Omega.$$

Note that by the $\mathbf{P}-$invariance of θ, we have $A \le B$ iff the family of inequalities

$$A(\theta^n \cdot) \le B(\theta^n \cdot), \quad n \in \mathbf{Z},$$

is valid $\mathbf{P}-$a.s. This implication of ordering for the entire translates of the random matrices will be tacitly taken for granted in all that follows. In this sense we may define for random matrices A with values ($\mathbf{P}-$a.s.) in $\mathcal{M}^-$ and B with values in $\mathcal{M}^+$ the set

$$\widehat{AB} = \{C : C \quad \text{random stochastic matrix} \quad \text{such that} \quad A \le C \le B\}.$$

Sets of this type are called *random interval stochastic matrices* with boundaries A and B. For any random ($\mathbf{P}-$a.s.) probability vector $p \in \Sigma_M$ and for any random interval stochastic matrix $\widehat{AB}$ we let $\widehat{AB}p$ be the set of random vectors

$$\{Cp : C \in \widehat{AB}\}.$$

These sets which have some relevance for numerical studies of invariant measures for chaotic dynamical systems (see Diamond et al. [11]) are described by the following theorem. To formulate it, for a pair of (random) matrices A, B with values in $\mathcal{M}^-$ resp. $\mathcal{M}^+$, and any subset I of $\{1, \cdots, M\}$ we define the (random) *(upper) (j,I)-flow* by

$$H_j(I, \widehat{AB}) = \min\{\sum_{i \in I} b_{ij}, 1 - \sum_{i \notin I} a_{ij}\}.$$

Theorem 3.1 *For a random vector $p \in \Sigma_M$ the set $\widehat{AB}p$ is described by the set of all random vectors $q \in \Sigma_M$ satisfying*

$$\sum_{j=1}^{M} p_j H_j(I, \widehat{AB}) \geq \sum_{i \in I} q_i(\theta\cdot) \quad \textit{for all} \quad I \subset \{1, \cdots, M\}.$$

Proof:
For p fixed, we let $Q(p)$ be the set of random matrices $q \in \Sigma_M$ satisfying

$$\sum_{j=1}^{M} p_j H_j(I, \widehat{AB}) \geq \sum_{i \in I} q_i(\theta\cdot) \quad \text{for all} \quad I \subset \{1, \cdots, M\}.$$

In these terms, we have to prove that $Q(p) = \widehat{AB}p$.

Step 1. Here we show the inclusion $\widehat{AB}p \cap \Sigma_M \subset Q(p)$. Let $q \in \widehat{AB}p \cap \Sigma_M$. Then there is a random matrix $C \in \widehat{AB}$ such that $Cp = q(\theta\cdot)$. This means that for any $I \subset \{1, \cdots, M\}$ we have

$$\sum_{i \in I} \sum_{j=1}^{M} c_{ij} p_j = \sum_{i \in I} q_i(\theta\cdot). \tag{6}$$

By definition of the ordering, we moreover have for any $1 \leq j \leq M, I \subset \{1, \cdots, M\}$

$$\sum_{i \in I} a_{ij} \leq \sum_{i \in I} c_{ij} \leq \sum_{i \in I} b_{ij}. \tag{7}$$

This implies the inequality

$$\sum_{i \in I} c_{ij} \leq \min\{\sum_{i \in I} b_{ij}, 1 - \sum_{i \notin I} a_{ij}\} = H_j(I, \widehat{AB}).$$

Hence, (6) and (7) together imply $q \in Q(p)$.

Step 2. Here we show that $Q(p) \subset \widehat{AB}p \cap \Sigma_M$. For this purpose, fix $q \in Q(p)$. For all $I \subset \{1, \cdots, M\}$ we then have

$$\sum_{j=1}^{M} p_j H_j(I, \widehat{AB}) \geq \sum_{i \in I} q_i(\theta\cdot). \tag{8}$$

We have to prove the existence of a random matrix $C \in \widehat{AB}$ satisfying $Cp = q(\theta\cdot)$. For each individual $\omega \in \Omega$ the existence of an appropriate matrix $C(\omega)$ is guaranteed by the deterministic result from Diamond et al. [12]. The problem is the measurability of the dependence on ω. This will be established by discretizing the problem in the following way. For each $k \in \mathbf{N}$, let us choose partitions $(A_n^k)_{n \in \mathbf{N}}$ of Σ_M and $(B_n^k)_{n \in \mathbf{N}}$ of S_M into sets of diameter at most $\frac{1}{k}$ with respect to the Euclidean distances on $\mathbf{R}^M$ and $\mathbf{R}^{M \times M}$. Assume in addition that the partition sequences are nested in k, i.e. each set of the sequence of degree $k+1$ is contained in a set of the sequence of degree k. For $n, m, r, s \in \mathbf{N}$, define

$$D_{nmrs}^k = \{\omega \in \Omega : A(\omega) \in B_n^k, B(\omega) \in B_m^k, p(\omega) \in A_r^k, q(\theta\omega) \in A_s^k\}.$$

Obviously, this yields a measurable partition, i.e. a family of pairwise disjoint measurable sets for which

$$\cup_{n,m,r,s\in\mathbf{N}} D^k_{nmrs} = \Omega$$

for any $k \in \mathbf{N}$. Next, for each $k, n, m, r, s \in \mathbf{N}$ for which this is possible, choose $\omega^k_{nmrs} \in D^k_{nmrs}$ and, using the deterministic result, choose $C^k_{nmrs} \in A(\omega^k_{nmrs})\widehat{B}(\omega^k_{nmrs})$ such that

$$C^k_{nmrs}\, p(\omega^k_{nmrs}) = q(\theta\omega^k_{nmrs}).$$

For each $k \in \mathbf{N}$, we set

$$C^k = \sum_{n,m,r,s\in\ n} C^k_{nmrs} 1_{D^k_{nmrs}},$$

and thus obtain a sequence of $\mathcal{S}_M-$valued random variables for which by definition

$$|C^k p - q(\theta\cdot)| \leq \frac{c}{k}$$

with a constant $c > 0$, which is bounded in $L^0(\Omega, \mathcal{F}, \mathbf{P})$, and at most a distance $\frac{c}{k}$ from the interval stochastic matrix $\widehat{AB}$. By boundedness, Föllmer and Schied [14], p. 37, allows us to choose a strictly increasing sequence $(\sigma_k)_{k\in\mathbf{N}}$ of random integers such that the sequence $(C^{\sigma_k})_{k\in\mathbf{N}}$ converges $\mathbf{P}-$a.s. to a random variable C with values in $\mathcal{S}_M$. This random variable obviously takes $\mathbf{P}-$a.s. values in $\widehat{AB}$, and fulfills the desired equation

$$Cp = q(\theta\cdot).$$

This completes the proof. $\diamond$

We call a random vector p with values in Σ_M *semi-invariant* for the random interval stochastic matrix $\widehat{AB}$ if for each $I \subset \{1, \cdots, M\}$ we have

$$\sum_{j=1}^{M} p_j H_j(I, \widehat{AB}) \geq \sum_{i\in I} p_i(\theta\cdot).$$

And finally, for any stochastic matrix C we denote by $\mathbf{Fix}(C)$ the set of all random vectors $x \in \Sigma_M$ for which $Cx = x(\theta\cdot)$. In these terms, we obtain the following corollary which will be important below.

Corollary 3.1 *The random set*

$$\bigcup_{C\in\widehat{AB}} \mathbf{Fix}(C)$$

consists of all random semi-invariant stochastic vectors in the random interval stochastic matrix $\widehat{AB}$.

4 Some results on the weak convergence of random measures

Another short section with preliminaries on the weak topology for random invariant measures is required. We consider the *Hausdorff semi-distance* between $A, B \in \mathcal{B}(X)$ given by

$$\widehat{\chi}(A, B) = \inf\{\epsilon > 0 : A \subset G_\epsilon(B)\},$$

and define the *Hausdorff distance* between $A_1, A_2 \in \mathcal{B}(X)$ as

$$\chi(A_1, A_2) = \max\{\widehat{\chi}(A_1, A_2), \widehat{\chi}(A_2, A_1)\}.$$

Let us first recall some basic facts concerning the weak topology for probability measures on metric spaces.

Lemma 4.1 *For probability measures $\mu_1, \mu_2, \nu_1, \nu_2$ on $\mathcal{B}(X)$ and for $\lambda \in [0, 1]$ we have*

$$\rho(\lambda\mu_1 + (1-\lambda)\mu_2, \lambda\nu_1 + (1-\lambda)\nu_2) \leq \max\{\rho(\mu_1, \nu_1), \rho(\mu_2, \nu_2)\}. \tag{9}$$

The straightforward proof is omitted.

Lemma 4.2 *Let the sequence $(\mu^{(n)})_{n\in\mathbf{N}}$ of random probability measures converge to the random probability measure μ. Moreover, suppose that the sequence of Borel sets $(B_k)_{k\in\mathbf{N}}$ converges in the Hausdorff distance to the closed set B. Then we have for $\mathbf{P}-$a.e. $\omega \in \Omega$*

$$\limsup_{k\to\infty} \mu^{(k)}_{\theta^n\omega}(B_k)) \leq \mu_{\theta^n\omega}(B), \quad n \in \mathbf{N}.$$

Proof:
By definition, $R(\mu^{(k)}, \mu) \to 0$ as $k \to \infty$, hence also $\rho(\mu^{(k)}_{\theta^n\cdot}, \mu_{\theta^n\cdot}) \to 0$ for all $n \in \mathbf{N}$ $\mathbf{P}-$a.s. as $k \to \infty$. Pick ω from a set of probability 1 on which this convergence holds true for the $\omega-$sections of the measures, and pick $n \in \mathbf{N}$. Then for any $\epsilon > 0$ the sequence $(B_k)_{k\in\mathbf{N}}$ is absorbed by the open set $G_\epsilon(B)$, and therefore we have

$$\mu_{\theta^n\omega}(G_\epsilon(B)) \geq \limsup_{k\to\infty} \mu^{(k)}_{\theta^n\omega}(B_k) - \epsilon.$$

Hence by the outer regularity of $\mu_{\theta^n\omega}$ and due to $S = \cap_{\epsilon>0} G_\epsilon(S)$, we obtain

$$\mu_{\theta^n\omega}(B) \geq \limsup_{k\to\infty} \mu^{(k)}_{\theta^n\omega}(B_k) - \epsilon,$$

ϵ being arbitrary, the desired inequality follows. $\diamond$

Now let $\overline{\mathrm{Gr}(f^*)}$ denote the weak closure of the graph $\mathrm{Gr}(f^*)$ with respect to the topology induced by R on the space $\mathcal{F} \otimes \mathcal{B}(X) \times \mathcal{F} \otimes \mathcal{B}(X)$.

Lemma 4.3 *Let $(\mu, \nu) \in \overline{Gr(f^*)}$. Then for $\mathbf{P}-$a.e. $\omega \in \Omega$ we have*

$$\nu_{\theta\omega}(B) \leq \mu_\omega(\overline{f^{-1}}(B)), \quad B \in \mathcal{B}(X).$$

Proof:
Choose sequences of random probability measures $\mu^{(n)}, \nu^{(n)}, n \in \mathbf{N}$, such that

$$\lim_{n\to\infty} R(\mu^{(n)}, \mu) = 0, \quad \lim_{n\to\infty} R(\nu^{(n)}, \nu) = 0, \tag{10}$$

and

$$\nu^{(n)} = f^* \mu^{(n)}, \quad n \in \mathbf{N}. \tag{11}$$

In terms of the factorizations this means for $\mathbf{P}-$a.e. $\omega \in \Omega$

$$\lim_{n\to\infty} \rho(\mu_\omega^{(n)}, \mu_\omega) = 0, \quad \lim_{n\to\infty} \rho(\nu_{\theta\omega}^{(n)}, \nu_{\theta\omega}) = 0, \tag{12}$$

and

$$\nu_{\theta\omega}^{(n)} = \mu_\omega(f^{-1}(\omega, \cdot)), \quad n \in \mathbf{N}. \tag{13}$$

Pick ω from a set of probability 1 for which (12) and (13) hold. Using the Portmanteau theorem, we deduce that for any $m \in \mathbf{N}$ there is $m \leq n(m) \in \mathbf{N}$ such that for $n \geq n(m)$ we have

$$\nu_{\theta\omega}(B) \leq \nu_{\theta\omega}^{(n)}(G_{\frac{1}{m}}(B)) + \frac{1}{m}, \quad B \in \mathcal{B}(X), \tag{14}$$

hence

$$\nu_{\theta\omega}(B) \leq \mu_\omega^{(n)}(f^{-1}(\omega, G_{\frac{1}{m}}(B))) + \frac{1}{m}, \quad B \in \mathcal{B}(X). \tag{15}$$

Letting $m \to \infty$, we find

$$\nu_{\theta\omega}(B) \leq \limsup_{m\to\infty} \mu_\omega^{(n(m))}(f^{-1}(\omega, G_{\frac{1}{m}}(B))), \quad B \in \mathcal{B}(X). \tag{16}$$

But

$$\lim_{m\to\infty} \chi(f^{-1}(\omega, G_{\frac{1}{m}}(B)), f^{-1}(\omega, B)) = 0. \tag{17}$$

Now we apply Lemma 4.2, which is justified due to (16), and (17) to obtain

$$\nu_{\theta\omega}(B) \leq \mu_\omega(f^{-1}(\omega, B)), \quad B \in \mathcal{B}(X).$$

This completes the proof. $\diamond$

Finally, let X_M be a finite subset of X, and consider random Markov chains p on X_M as explained above. For any random probability measure μ on the product of $\mathcal{F}$ and the power set of X_M we define

$$p\mu_i(\theta\cdot) = \sum_{j=1}^{M} p_{ij}\mu_j.$$

In particular, for the Dirac measure δ_k at k we obtain

$$(p\delta_k)_i(\theta\cdot) = p_{ki}, \quad 1 \leq k \leq M.$$

For a set Q of pairs of random probability measures on $\Omega \times X$ and another pair (μ, ν) we define

$$\rho((\mu, \nu), Q) = \inf_{(\mu', \nu') \in Q} \max\{\rho(\mu, \mu'), \rho(\nu, \nu')\}.$$

With these preliminaries, we have the following inequality.

Lemma 4.4 *Let p be a random Markov chain. For any $1 \leq k \leq M$ we have*

$$\rho((\delta_k, p\delta_k), Gr(f^*)) \leq [D(p,f)]^{\frac{1}{2}} \quad \mathbf{P}-a.s.$$

Proof:
Fix $\omega \in \Omega$ such that $D(p(\omega), f(\omega,\cdot)) < \infty$. Let γ be a real number such that $\gamma > [D(p(\omega), f(\omega,\cdot)]^{\frac{1}{2}}$. Our task is to construct a measure μ_ω on $\mathcal{B}(X)$ such that

$$\rho(\mu, \delta_k) \leq \gamma, \tag{18}$$

and

$$\rho(f(\omega,\cdot)^*(\mu_\omega), p(\omega)\delta_k) \leq \gamma. \tag{19}$$

This being granted we may finally construct a measurable version of $\omega \mapsto \mu_\omega$ in a way similar to step 2 of the proof of Theorem 3.1. Let

$$C(k) = \{1 \leq i \leq M : \mathrm{dist}_{X\times X}((k,i), \mathrm{Gr}(f(\omega,\cdot)) \geq \gamma\}.$$

Then Chebyshev's inequality gives for $1 \leq k \leq M$

$$p(\omega)\delta_k(C(k)) \leq \gamma. \tag{20}$$

Consider the map $\Phi : \{1, \cdots, M\} \to X$ which satisfies $\Phi(i) = k$ if $i \in C(k)$, and

$$d(\Phi(i), k) \leq \gamma \quad \text{and} \quad d(f(\omega, \Phi(i)), i) \leq \gamma.$$

Then by definition for all $1 \leq i \leq M$

$$d(\Phi(i), k) \leq \gamma. \tag{21}$$

We now define

$$\mu_{\theta\omega} = \sum_{i=1}^{M} p(\omega)\delta_k(\{i\})\delta_{\Phi(i)}.$$

Then (18) is a consequence of (21), whereas (19) follows from (20). $\diamond$

Lemma 4.5 *Let μ be a random $M-$vector. Then we have $\mathbf{P}-a.s.$*

$$\rho((\mu, (p\mu)_{\theta\cdot}), Gr(f^*)) \leq [D(p,f)]^{\frac{1}{2}}.$$

Proof:
This is a simple consequence of Lemma 4.1, Lemma 4.4 and the linearity of f^*, with arguments for fixed ω in a set of measure 1. $\diamond$

5 Semi-invariant and approachable measures

We now fix a random mapping f generating the random dynamical system ϕ. We call a probability measure μ on $\mathcal{F} \otimes \mathcal{B}(X_{M_k})$ *approachable* if there exists an approximating sequence $(p^{(k)})_{k \in \mathbf{N}}$ of random Markov chains for ϕ and a sequence of random measures $\mu^{(k)}$ on $\mathcal{F} \otimes \mathcal{B}(X_{M_k})$ invariant for $p^{(k)}$ such that $R(\mu^{(k)}, \mu) \to 0$. Speaking about approximating chains in the sequel, we shall always suppose that $p^{(k)}$ is associated with a discretization X_{M_k} of X of maximal step size h_{M_k} and we suppose $h_{M_k} \to 0$ as $k \to \infty$.

Our aim is to show that any invariant measure of ϕ is approachable. As was observed in [12] in the deterministic setting, approachability is equivalent to a notion of semi-invariance which is more general than invariance. We will extend this notion to the setting of random measures, and prove that it is equivalent to approachability in the sense above.

For a mapping $g : X \to X$ let $\overline{\mathrm{Gr}(g)}$ be the closure of the set $\mathrm{Gr}(g)$. For any set $B \in \mathcal{B}(X)$ take $\overline{g^{-1}}(B)$ to be defined by

$$\overline{g^{-1}}(B) = \{x \in X : \quad \text{there is} \quad y \in \overline{B} \quad \text{such that} \quad (x, y) \in \overline{\mathrm{Gr}(g)}\}.$$

We call a random probability measure μ with disintegration $\mu_\cdot$ $\pi-$semi-invariant if $\mathbf{P}-$a.s. the following statement holds:

$$\mu_{\theta \cdot}(B) \leq \mu_\cdot(\overline{f^{-1}}(\cdot, B))$$

for all $B \in \mathcal{B}(X)$. (Recall that π is the skew product flow generated by the random difference equation 1). Via a standard measure theoretic argument involving the uniqueness of measures defined on a countable generator which is stable for intersections, this statement is equivalent with

$$\mu_{\theta^{n+1} \cdot}(B) \leq \mu_{\theta^n \cdot}(\overline{f^{-1}}(\cdot, B)) \quad \text{for any} \quad B \in \mathcal{B}(X), n \in \mathbf{Z},$$

$\mathbf{P}-$a.s. Of course, every invariant measure for π is $\pi-$semi-invariant.

5.1 Properties of semi-invariant measures

We shall now study the property of semi-invariance more closely. We first show that it is genuine to non-continuous generators f. Indeed, if for $\omega \in \Omega$, $f(\omega, \cdot)$ is continuous, then any semi-invariant measure is invariant.

Theorem 5.1 *Suppose that μ is $\pi-$semi-invariant, and that for $\mathbf{P}-$a.e. $\omega \in \Omega$ the mapping $f(\omega, \cdot)$ is continuous $\mu_\omega-$a.e. in X. Then μ is $\pi-$invariant.*

Proof:
Suppose that there exists a set of positive measure C such that for $\omega \in C$ the inequality

$$\mu_{\theta^{n+1}\omega}(B) \geq \mu_{\theta^n \omega}(\overline{f^{-1}}(\theta^n \omega, B))$$

is not valid for all $n \in \mathbf{N}, B \in \mathcal{B}(X)$. By eventually removing a set of measure 0, we may assume that for any $\omega \in C$ the mappings $f(\theta^n\omega, \cdot)$ are continuous $\mu_{\theta^n\omega}-$a.e., say, outside a null set $N_{\theta^n\omega}$, and that

$$\mu_{\theta^{n+1}\omega}(B) \leq \mu_{\theta^n\omega}(\overline{f^{-1}}(\theta^n\omega, B)), \quad n \in \mathbf{N}, B \in \mathcal{B}(X). \tag{22}$$

Choose $\omega \in C$, to find $n \in \mathbf{N}$ and $B \in \mathcal{B}(X)$ for which we have

$$\mu_{\theta^{n+1}\omega}(B) < \mu_{\theta^n\omega}(\overline{f^{-1}}(\theta^n\omega, B)).$$

To abbreviate, write

$$\nu_{n+1} = \mu_{\theta^{n+1}\omega}, \quad \nu_n = \mu_{\theta^n\omega}, \quad g_n = f(\theta^n\omega, \cdot).$$

Choose a closed set F such that

$$B \cap F = \emptyset \quad \text{and} \quad \nu_{n+1}(F) + \nu_n(\overline{g_n^{-1}}(B)) > 1. \tag{23}$$

The first part of (23) implies that

$$\overline{g_n^{-1}}(B) \cap \overline{g_n^{-1}}(F) \subset D(g_n), \tag{24}$$

where $D(g_n)$ denotes the set of discontinuity points of g_n. By ν_n-a.e. continuity of g_n, we get

$$\nu_n(\overline{g_n}^{-1}(B) \cap \overline{g_n^{-1}}(F)) = 0.$$

By (22), we have on the other hand $\nu_n(\overline{g_n^{-1}}(C)) \geq \nu_n(C)$. So the second part of (23) implies

$$\nu_n(\overline{g_n^{-1}}(B) \cup \overline{g_n^{-1}}(C)) = \nu_n(\overline{g_n^{-1}}(B)) + \nu_n(\overline{g_n^{-1}}(C)) > 1.$$

But this is not possible because ν_n is a probability measure. $\diamond$

If the random generator f is discontinuous on sets of positive $\mu_\omega-$measure, for a relevant set of ω, semi-invariant measures are more significant than invariant ones. For example, the set of semi-invariant measures is much more robust with respect to perturbations of f.

Let us next define a multi-valued operator on the space of random probability measures on $\mathcal{F} \otimes \mathcal{B}(X)$. For a random probability measure μ we set

$$\begin{aligned} F^*\mu \;\; = \;\; & \{\nu : \nu \,\text{random probability measure, } \nu_{\theta^{n+1}\omega}(B) \leq \mu_{\theta^n\omega}(\overline{f^{-1}}(\theta^n\omega, B)) \\ & \text{for all } B \in \mathcal{B}(X), n \in \mathbf{Z}, \mathbf{P} - \text{a.a } \omega \in \Omega\}. \end{aligned}$$

Moreover, denote by $\overline{f^*}$ the (possibly also multi-valued) weak closure of the operator f^*, i.e. its closure with respect to the metric R. For a random probability measure μ the set $\overline{f^*}\mu$ is just given by the set of all weak limits of sequences $(f^*\mu^{(n)})_{n\in\mathbf{N}}$ for which the sequence $(\mu^{(n)})_{n\in\mathbf{N}}$ converges in the $R-$metric to μ.

Theorem 5.2 *The operators F^* and $\overline{f^*}$ coincide.*

Proof:
We have to prove that

$$\overline{\mathrm{Gr}(f^*)} = \mathrm{Gr}(F^*). \tag{25}$$

From Lemma (4.3) we have

$$\overline{\mathrm{Gr}(f^*)} \subset \mathrm{Gr}(F^*). \tag{26}$$

To prove the reverse inclusion, we introduce the multi-valued operator G^* which maps random probability measures μ to the collection of weak limits of sequences of the type $(p_{M_k}\mu^{(k)})_{k\in\mathbf{N}}$ with an approximating sequence $(p_{M_k})_{k\in\mathbf{N}}$ of random Markov chains for f, defined as random $M_k\times M_k-$matrices, and a sequence $(\mu^{(k)})_{k\in\mathbf{N}}$ of random M_k-vectors converging weakly to μ. By virtue of Lemma 4.5, we have for $\mathbf{P}-$a.e. $\omega\in\Omega$

$$\mathrm{Gr}(G^*_\omega) \subset \overline{\mathrm{Gr}(f^*_\omega)}. \tag{27}$$

Due to dominated convergence, (27) implies

$$\mathrm{Gr}(G^*) \subset \overline{\mathrm{Gr}(f^*)}. \tag{28}$$

To obtain the reverse inclusion, it therefore remains to see that

$$\mathrm{Gr}(G^*) \subset \mathrm{Gr}(F^*). \tag{29}$$

This part is almost identical to the second part of the proof of the subsequent main result. For details we refer to that proof. $\diamond$

5.2 Semi-invariant measures are approachable

We are ready to state the main result of this paper.

Theorem 5.3 *A probability measure μ on $\Omega\times X$ is $\pi-$approachable through a sequence of random Markov chains if and only if it is $\pi-$semi-invariant.*

Proof:
The proof consists of two steps. In the first one, we shall show that any approachable measure is semi-invariant. In the more complicated second step, we shall show that every semi-invariant measure is approachable. We recall that inequalities between random variables are in fact $\mathbf{P}-$a.s. inequalities.

Step 1: Fix a random probability measure μ which is approachable. We shall prove that for $\mathbf{P}-$a.e. $\omega\in\Omega$ the inequality

$$\mu_{\theta\omega}(B) \le \mu_\omega(\overline{f^{-1}}(\omega, B)), \quad B\in\mathcal{B}(X)$$

holds true. By applying Lemma 4.3, we reduce this statement to the following statement being true for $\mathbf{P}-$a.e. $\omega\in\Omega$

$$(\mu_\omega, \mu_{\theta\omega}) \in \overline{\mathrm{Gr}(f^*)(\omega,\cdot)}. \tag{30}$$

Now choose an approximating sequence $(p^{(k)})_{k\in\mathbf{N}}$ of Markov chains and a sequence of random measures $(\mu^{(k)})_{k\in\mathbf{N}}$ such that $\mu^{(k)}$ is invariant for $p^{(k)}$ for all $k\in\mathbf{N}$ and such that $R(\mu^{(k)},\mu)\to 0$, hence $\rho(\mu^{(k)}_\omega,\mu_\omega)\to 0$ for $\mathbf{P}-$a.e. $\omega\in\Omega$. Now note that for $\mathbf{P}-$a.e. $\omega\in\Omega$ we have $p^{(k)}(\omega)\mu^{(k)}_\omega=\mu_{\theta\omega}, k\in\mathbf{N}$. Hence Lemma 4.5 yields for $\mathbf{P}-$a.e. $\omega\in\Omega$

$$\rho((\mu^{(k)}_{\theta\omega},\mu^{(k)}_\omega),\mathrm{Gr}(f^*)_\omega)=\rho((\mu^{(k)}_{\theta\omega},(p^{(k)}\mu^{(k)})_\omega),\mathrm{Gr}(f^*)_\omega)\le[D(p^{(k)}(\omega),f(\omega,\cdot))]^{\frac{1}{2}}. \quad (31)$$

Since $(p^{(k)})_{k\in\mathbf{N}}$ is approximating, we deduce

$$\rho((\mu_{\theta\omega},\mu_\omega),\mathrm{Gr}(f^*))=0, \quad (32)$$

which implies (30).

Step 2: Assume now that the random probability measure μ is semi-invariant, that is for $\mathbf{P}-$a.e. $\omega\in\Omega$ the family of inequalities

$$\mu_{\theta\omega}(B)\le\mu_\omega(\overline{f^{-1}}(\omega,B)),\quad B\in\mathcal{B}(X) \quad (33)$$

is satisfied. Fix $k\in\mathbf{N}$. Then approachability will be a consequence of the following Lemma. Once this Lemma is proved, Step 2 will be accomplished and therefore the proof of the Theorem finished.

Lemma 5.1 *Let $\gamma>1$. There exists a random Markov chain $p^{(k)}$ and a random probability measure $\mu^{(k)}$ on X_{M_k} invariant for $p^{(k)}$ and such that*

$$D(p^{(k)},f)\le\gamma h_{M_k}, \quad (34)$$

$$R(\mu^{(k)},\mu)\le h_{M_k}. \quad (35)$$

Proof:
We first construct the measure $\mu^{(k)}$. To this end, abbreviate $d=M_k$, and enumerate X_{M_k} by $x_1,\cdots,x_d$. Let $(U(i):1\le i\le d)$ be a partition of X into Borel sets satisfying

$$\begin{aligned}&(\mathbf{P1})\quad \text{any element } x_i\in X_{M_k} \text{ belongs only to the set } U(i),\ \text{for all } 1\le i\le d,\\ &(\mathbf{P2})\quad U(i)\subset\{x\in X: d(x,x_i)\le h_{M_k}\} \text{ for all } 1\le i\le d.\end{aligned} \quad (36)$$

Now define

$$\mu^{(k)}(\omega)(i)=\mu(\omega,U(i)),\quad 1\le i\le d. \quad (37)$$

Evidently, the random vector $\mu^{(k)}$ is a probability vector. If we interpret this vector as a random linear combination of Dirac measures charging the points of X_{M_k} to measure distances between $\mu^{(k)}$ and μ, we have by construction

$$R(\mu^{(k)},\mu)\le h_{M_k}.$$

It remains to construct the random chain $p^{(k)}$ for which $\mu^{(k)}$ is invariant and which satisfies (34). In fact we shall construct a chain which does better than just satisfy (34). Let $C_k(\gamma)$ be the collection of all random $d\times d-$matrices p satisfying

$$p_{ij}=0\quad\text{if}\quad \mathrm{dist}((x_i,x_j),\mathrm{Gr}(f))>\gamma h_{M_k}.$$

Note that $p \in C_k(\gamma)$ implies (34). We therefore have to construct a random chain $p^{(k)} \in C_k(\gamma)$ with invariant random vector $\mu^{(k)}$. At this point random interval stochastic matrices enter the game. It is evident that

$$C_k(\gamma) = \widehat{AB}, \tag{38}$$

where $A = 0$ and B is the random $d \times d-$matrix defined by

$$b_{ij}(\omega) = \begin{cases} 1 & \text{if } \mathrm{dist}((x_i, x_j), \mathrm{Gr}(f(\omega, \cdot)) \leq \gamma h_{M_k}, \\ 0 & \text{otherwise.} \end{cases} \tag{39}$$

According to Theorem 3.1 and its Corollary, we have to show that $\mu^{(k)}$ is a semi-invariant random vector for the random interval stochastic matrix $\widehat{AB}$. This means that for any $I \subset \{1, \cdots, d\}$ we have to establish

$$\sum_{j=1}^{d} \mu^{(k)}(j) H_j(I, \widehat{AB}) \geq \sum_{j \in I} \mu^{(k)}(j). \tag{40}$$

In our particular case, the upper $(j, I)-$flow H_j is given by the simple relationship

$$H_j(I, \widehat{AB}) = \begin{cases} 1 & \text{if there is } i \in I \text{ such that } b_{ij} = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{41}$$

By definition of the b_{ij} and (41), the inequality to establish can be rephrased in the form

$$\sum_{j \in I} \mu_{\theta \cdot}^{(k)}(j) \leq \sum_{j \in J(I)} \mu^{(k)}(j), \tag{42}$$

where

$$J(I) = \{l : 1 \leq j \leq d, \text{ there is } i \in I \text{ such that } \mathrm{dist}((x_i, x_j), \mathrm{Gr}(f)) \leq \gamma h_k\}.$$

The measurability of the random integer set $J(I)$ is due to the measurability of f. Recalling the definition of $\mu^{(k)}$, the inequality (42) takes the form

$$\mu_{\theta \cdot}(\cup_{i \in I} U(i)) = \sum_{i \in I} \mu_{\cdot}(U(i)) \leq \sum_{j \in J(I)} \mu_{\cdot}(U(j)). \tag{43}$$

Take $\epsilon = \gamma - 1$. Then by semi-invariance

$$\mu_{\theta \cdot}(\cup_{i \in I} U(i)) \leq \mu_{\cdot}(G_\epsilon[\cup_{i \in I} f^{-1}(\cdot, G_\epsilon(U(i)))]. \tag{44}$$

Let $B_1 = \cup_{i \in I} f^{-1}(\cdot, G_\epsilon(U(i)))$. Then (43) will be proved, once we have established

$$\mu_{\cdot}(B_1) \leq \sum_{j \in J(I)} \mu_{\cdot}(U(j)). \tag{45}$$

To do this, recall our construction of $\mu^{(k)}$ due to which

$$B_1 \subset G_\epsilon[\cup_{i \in I} \{x \in X : d(f(\cdot, x), x_i) \leq \gamma h_k\}]. \tag{46}$$

Now suppose that $U(j) \cap B_1 \neq \emptyset$. Then by construction there is an $i \in I$ such that $\mathrm{dist}((x_i, x_j), \mathrm{Gr}(f)) \leq \gamma h_{M_k}$. This just means $B_1 \subset \cup_{j \in J(I)} U(j)$, which in turn implies the desired

$$\mu_{\cdot}(B_1) \leq \sum_{j \in J(I)} \mu_{\cdot}(U(j)).$$

This completes the proof of Lemma 5.1. $\diamond$

References

[1] L. Arnold, *Random Dynamical Systems.* Springer: Berlin 1998.

[2] T. Bogenschütz, *Equilibrium States for Random Dynamical Systems.* Doctoral dissertation, Universität Bremen 1993.

[3] T. Bogenschütz, V.M. Gundlach, *Symbolic dynamics for expanding random dynamical systems.* Random and Computational Dynamics 1 (1992), 219-227.

[4] T. Bogenschütz, V.M. Gundlach, *Ruelle's transfer operator for random subshifts of finite type.* Ergod. Th. Dynam. Sys. 15 (1995), 413-447.

[5] M. Dellnitz, G. Froyland and O. Junge, The algorithms behind GAIO – set oriented numerical methods for dynamical systems, in *Ergodic Theory; Analysis and Efficient Simulation of Dynamical Systems.* Editor: B. Fiedler, Springer–Verlag, 2001, pp. 144–174.

[6] M. Dellnitz and O. Junge, On the approximation of complicated dynamical behaviour. *SIAM J. Numer. Anal.* **36** (1998), 491-515.

[7] M. Denker, *Finite generators for ergodic measure-preserving transformations.* Z. Wahrscheinlichkeitstheorie verw. Gebiete 29 (1974), 45-55.

[8] M. Denker, C. Grillenberger and K. Sigmund, *Ergodic Theory on Compact Spaces.* Springer LNM 527, Springer: Berlin 1976.

[9] M. Denker, M. Gordin and St. M. Heinemann, On the relative variational principle for fiber expanding mappings. *Ergodic Theory Dynamical Systems* **22** (2002), 757–782.

[10] P. Diamond and P. E. Kloeden, Spatial discretization of mappings, *J. Computers Math. Applns.* **26** (1993), 85-94.

[11] P. Diamond, P. Kloeden and A. Pokrovskii, An invariant measure arising in computer simulation of a chaotic dynamical system. *J. Nonlinear Sciences* **4** (1994), 59-68.

[12] P. Diamond, P. Kloeden and A. Pokrovskii, Interval Stochastic Matrices: A Combinatorial Lemma and the Computation of Invariant Measures of Dynamical Systems. *J. Dynamics Differential Equations* **7** (1995), 341-364.

[13] S. Ethier and T. Kurtz, *Markov Processes: Characterization and Convergence.* Wiley: New York 1986.

[14] H. Föllmer and A. Schied, *Stochastic Finance: An Introduction in Discrete Time.* Walter de Gruyter: Berlin 2002.

[15] Ch. Grillenberger, U. Krengel, *On marginal distributions and isomorphisms of stationary processes.* Math. Z. 149 (1976), 131-154.

[16] V.M. Gundlach and O. Steinkamp, Products of random rectangular matrices. *Math. Nachrichten* **212** (2000), 51-76.

[17] R. Guder, M. Dellnitz and E. Kreuzer, Numerical approximation of invariant measures and error estimates. *Chaos Fractals Solitons* **8** (1997), 525-534.

[18] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems.* Cambridge University Press: Cambridge 1995.

[19] W. Krieger, *On entropy and generators of measure preserving transformations.* Trans. Amer. Math. Soc. 149 (1970), 453-464.

[20] K. Mischaikow, Topological techniques for efficient rigorous computation in dynamics. *Acta Numerica* (2002), pp. 435-477.

[21] D. Ruelle, *Thermodynamic formalism.* Encyclopedia of Mathematics and itas Applications. Volume 5. Addison-Wesley: Reading, MA (1978).

[22] K. Sigmund, Ergodic averages for axiom A diffeomorphisms. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **20** (1971), 319–324.

[23] A.M. Stuart and A.R. Humphries, *Numerical Analysis and Dynamical Systems,* Cambridge University Press: Cambridge 1996.