Computing the speed of convergence of ergodic averages and pseudorandom points in computable dynamical systems

Stefano Galatolo, Mathieu Hoyrup, Cristobal Rojas

HAL Id: inria-00517372

https://hal.inria.fr/inria-00517372

Submitted on 14 Sep 2010

Computing the speed of convergence of ergodic averages and pseudorandom points in computable dynamical systems

Stefano Galatolo

Dipartiento di matematica applicata
Universita di Pisa

s.galatolo@ing.unipi.it

Mathieu Hoyrup

LORIA, Vandoeuvre-l es-Nancy, France

Mathieu.Hoyrup@loria.fr

Cristóbal Rojas

Fields Institute, Toronto, Canada

cristobal.rojas@utoronto.ca

A pseudorandom point in an ergodic dynamical system over a computable metric space is a point which is computable but its dynamics has the same statistical behavior as a typical point of the system.

It was proved in [2] that in a system whose dynamics is computable the ergodic averages of computable observables converge effectively. We give an alternative, simpler proof of this result.

This implies that if also the invariant measure is computable then the pseudorandom points are a set which is dense (hence nonempty) on the support of the invariant measure.

1 Introduction

We will consider abstract algorithmic questions concerning the evolution of a dynamical system. In particular, the algorithmic estimation of the speed of convergence of ergodic averages and the recursive construction of points whose dynamics is typical for the system.

This latter problem is related to the possibility of computer simulations, as actual computers can only calculate the evolution of computable initial conditions.

Let X be a separable metric space and let the iterations of a map $T : X \to X$ define a dynamics. Let μ be an invariant measure for the dynamics ($\mu(A) = \mu(T^{-1}(A))$ for each measurable set A). The famous Birkhoff ergodic theorem says that if μ is ergodic (the only sets which are invariant for the dynamics have full or null measure) then

$$\lim_{n \to \infty} \frac{1}{n} \sum f(T^n(x)) = \int f \, d\mu \, , \, \mu - \text{almost everywhere}.$$

Similar results can be obtained for the convergence in the L^2 norm, and others. The above result tells that, if the system is ergodic, there is a fulll measure set of points for which the averages of the values of the observable f along its trajectory (time averages) coincides with the spatial average of the observable f. Such a point could be called *typical* for f. Points which are typical for each f which is continuous with compact support are called typical for the measure μ (see Definition 12).

Estimating the speed of convergence. Many, more refined results are linked to the speed of convergence of the above limit. We consider this problem from the point of view of Computable Analysis. In the paper [2] some abstract results imply that in a computable ergodic dynamical system, the speed of convergence of such averages can be algorithmically estimated. On the other hand it is also shown that there are non

ergodic systems where this kind of estimations are not possible. In this paper we present a simple, direct way to prove the following result:

Theorem A. *If* (X, μ, T) *is an ergodic dynamical system and T is a.e. computable, then for each computable L^1 observable f the ergodic average $A_n(x) = \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x))$ converge effectively a.e. to $\int f \, d\mu$ (see Definition 11 for the precise definition of this kind of convergence).*

This theorem has some interesting consequences, as we are going to illustrate.

Computable points having typical statistical behavior. The set of computable points in X (see Definition 3) being countable, is a very small (invariant) set, compared to the whole space. For this reason, a computable point could be rarely be expected to be typical for the dynamics, as defined before. More precisely, the Birkhoff ergodic theorem and other theorems which hold for a full measure set, cannot help to decide if there exist a computable point which is typical for the dynamics. Nevertheless computable points are the only points we can use when we perform a simulation or some explicit computation on a computer.

A number of theoretical questions arise naturally from all these facts. Due to the importance of forecasting and simulation of a dynamical system's behavior, these questions also have some practical motivation.

Problem 1 *Since simulations can only start with computable initial conditions, given some typical statistical behavior of a dynamical system, is there some computable initial condition realizing this behavior? how to choose such points?*

Such points could be called *pseudorandom* points. Meaningful simulations, showing typical behaviors of the dynamics can be performed if computable, pseudorandom initial conditions exist (and can be computed from the description of the system).

We remark that it is widely believed that computer simulations produce correct statistical behavior. The evidence is mostly heuristic. Most arguments are based on the various "shadowing" results (see e.g. [9] chapter 18). In this kind of approach (different from ours), it is possible to prove that in a suitable system every pseudo-trajectory, as the ones which are obtained in simulations with some computation error, is close to a real trajectory of the system. However, even if we know that what we see in a simulation is near to some real trajectory, we do not know if this real trajectory is typical in some sense.

The main limit of this approach is that shadowing results hold only in particular systems, having some uniform hyperbolicity, while many physically interesting systems are not like this.

In our approach we consider real trajectories instead of "pseudo" ones and we ask if there is some computable point which behaves as a typical point of the space. Thanks to a kind of effective Borel-Cantelli lemma, in [6] the above problem is solved affirmatively for a class of systems which satisfies certain technical assumptions which includes systems whose decay of correlation is faster that $C \log^2(time)$. In this paper we prove the following more general result, as a consequence of the above Theorem A:

Theorem B. *If* (X, μ, T) *is a computable dynamical system as above and μ is a computable invariant ergodic measure, there exist computable points x for which it holds:*

$$\lim_{n \to \infty} \frac{f(x) + f(T(x)) + \ldots + f(T^{n-1}(x))}{n} = \int f \, d\mu \tag{1.1}$$

for any continuous function $f : X \to \mathbb{R}$ with compact support.

The above theorem hence states that in such systems there are pseudorandom points.

Physical measures and computability. To apply the above corollary to concrete systems the main difficulty is to verify that the invariant measure is computable. In [6] and [8] it is shown that this is verified for the physical[1] invariant measure (the natural invariant measure to be considered in this cases) in several classes of interesting systems as uniformly hyperbolic systems, piecewise expanding maps and interval maps with an indifferent fixed point. On the other hand there are computable systems having no computable invariant measure ([8]) this shows some subtlety in this kind of questions.

The way we handle computability on continuous spaces is largely inspired by representation theory (see [16],[3]). However, the main goal of that theory is to study, in general topological spaces, the way computability notions depend on the chosen representation. Since we focus only on computable metric spaces, we do not use representation theory in its general setting but instead present computability notions in a self-contained way, and hopefully accessible to non-specialists.

2 Computability

The starting point of recursion theory was to give a mathematical definition making precise the intuitive notions of *algorithmic* or *effective procedure* on symbolic objects. Several different formalizations have been independently proposed (by Church, Kleene, Turing, Post, Markov...) in the 30's, and have proved to be equivalent: they compute the same functions from $\mathbb{N}$ to $\mathbb{N}$. This class of functions is now called the class of *recursive functions*. As an algorithm is allowed to run forever on an input, these functions may be *partial*, i.e. not defined everywhere. The *domain* of a recursive function is the set of inputs on which the algorithm eventually halts. A recursive function whose domain is $\mathbb{N}$ is said to be *total*.

We now recall an important concept from recursion theory. A set $E \subseteq \mathbb{N}$ is said to be ***recursively enumerable (r.e.)*** if there is a (partial or total) recursive function $\varphi : \mathbb{N} \to \mathbb{N}$ enumerating E, that is $E = \{\varphi(n) : n \in \mathbb{N}\}$. If $E \neq \emptyset$, φ can be effectively converted into a total recursive function ψ which enumerates the same set E.

2.1 Algorithms and uniform algorithms

Strictly speaking, recursive functions only work on natural numbers, but this can be extended to the objects (thought as "finite" objects) of any countable set, once a numbering of its elements has been chosen. We will use the word *algorithm* instead of *recursive function* when the inputs or outputs are interpreted as finite objects. The operative power of algorithms on the objects of such a numbered set obviously depends on what can be effectively recovered from their numbers.

---

[1] In general, given (X,T) there could be infinitely many invariant probability measures. Among this class of measures, some of them are particularly important because they are related to what can be seen in real experiments. Suppose that we observe the behavior of the system (X,T) through a class of continuous functions $f_i : X \to \mathbb{R}$. We are interested in the statistical behavior of f_i along typical orbits of the system. Let us suppose that the time average along the orbit of x exists

$$A_x(f_i) = \lim_{n \to \infty} \frac{1}{n} \sum f_i(T^n(x)).$$

This is a real number for each f_i. Moreover $A_x(f_i)$ is linear and continuous with respect to small changes of f_i in the sup norm. Then the orbit of x acts as a measure μ_x and $A_x(f_i) = \int f_i \, d\mu_x$ (moreover this measure is also invariant for T). This measure is physically interesting if it is given by a "large" set of initial conditions. This set will be called the basin of the measure. If X is a manifold, it is said that an invariant measure is *physical* if its basin has positive Lebesgue measure (see [17] for a survey and more precise definitions).

More precisely, let X and Y be numbered sets such that the numbering of X is injective (it is then a bijection between $\mathbb{N}$ and X). Then any *recursive function* $\varphi : \mathbb{N} \to \mathbb{N}$ induces an *algorithm* $\mathscr{A} : X \to Y$. The particular case $X = \mathbb{N}$ will be much used.

For instance, the set $\mathbb{Q}$ of rational numbers can be injectively numbered $\mathbb{Q} = \{q_0, q_1, \ldots\}$ in an *effective* way: the number i of a rational a/b can be computed from a and b, and vice versa. We fix such a numbering: from now and beyond the rational number with number i will be denoted by q_i.

Now, let us consider computability notions on the set $\mathbb{R}$ of real numbers, introduced by Turing in [15].

Definition 1 *Let x be a real number. We say that:*

- *x is **lower semi-computable** if the set $\{i \in \mathbb{N} : q_i < x\}$ is r.e.*

- *x is **upper semi-computable** if the set $\{i \in \mathbb{N} : q_i > x\}$ is r.e.*

- *x is **computable** if it is lower and upper semi-computable.*

Equivalently, a real number is computable if and only if there exists an algorithmic enumeration of a sequence of rational numbers converging exponentially fast to x. That is:

Proposition 1 *A real number is **computable** if there is an algorithm $\mathscr{A} : \mathbb{N} \to \mathbb{Q}$ such that $|\mathscr{A}(n) - x| \leq 2^{-n}$ for all n.*

Uniformity. Algorithms can be used to define computability notions on many classes of mathematical objects. The precise definitions will be particular to each class of objects, but they will always follow the following scheme:

<div align="center">

An object O is ***computable*** if there is an algorithm

$$\mathscr{A} : X \to Y$$

which computes O in some way.

</div>

Each computability notion comes with a uniform version. Let $(O_i)_{i \in \mathbb{N}}$ be a sequence of computable objects:

<div align="center">

O_i is computable ***uniformly in i*** if there is an algorithm

$$\mathscr{A} : \mathbb{N} \times X \to Y$$

such that for all i, $\mathscr{A}_i := \mathscr{A}(i, \cdot) : X \to Y$ computes O_i.

</div>

For instance, the elements of a sequence of real numbers $(x_i)_{i \in \mathbb{N}}$ are uniformly computable if there is a algorithm $\mathscr{A} : \mathbb{N} \times \mathbb{N} \to \mathbb{Q}$ such that $|\mathscr{A}(i, n) - x_i| \leq 2^{-n}$ for all i, n.

In each particular case, the computability notion may take a particular name: computable, recursive, effective, r.e., etc. so the term "computable" used above shall be replaced.

2.2 Computable metric spaces

A computable metric space is a metric space with an additional structure allowing to interpret input and output of algorithms as points of the metric space. This is done in the following way: there is a dense subset (called ideal points) such that each point of the set is identified with a natural number. The choice of this set is compatible with the metric, in the sense that the distance between two such points is computable up to any precision by an algorithm getting the names of the points as input. Using these simple assumptions many constructions on metric spaces can be implemented by algorithms.

Definition 2 *A **computable metric space** (CMS) is a triple $\mathscr{X} = (X, d, S)$, where*

   (i) (X, d) *is a separable metric space.*

   (ii) $S = \{s_i\}_{i \in \mathbb{N}}$ *is a dense, numbered, subset of X called the set of **ideal points**.*

   (iii) *The distances between ideal points $d(s_i, s_j)$ are all computable, uniformly in i, j (there is an algorithm $\mathscr{A} : \mathbb{N}^3 \to \mathbb{Q}$ such that $|\mathscr{A}(i, j, n) - d(s_i, s_j)| < 2^{-n}$).*

    S is a numbered set, and the information that can be recovered from the numbers of ideal points is their mutual distances. Without loss of generality, we will suppose the numbering of *S* to be injective: it can always be made injective in an effective way.

    We say that in a metric space (X, d), a sequence of points $(x_n)_{n \in \mathbb{N}}$ converges *recursively* to a point *x* if there is an algorithm $D : \mathbb{Q} \to \mathbb{N}$ such that $d(x_n, x) \leq \varepsilon$ for all $n \geq D(\varepsilon)$.

Definition 3 *A point $x \in X$ is said to be **computable** if there is an algorithm $\mathscr{A} : \mathbb{N} \to S$ such that $(\mathscr{A}(n))_{n \in \mathbb{N}}$ converges recursively to x.*

    We define the set of ***ideal balls*** to be $\mathscr{B} := \{B(s_i, q_j) : s_i \in S, 0 < q_j \in \mathbb{Q}\}$ where $B(x, r) = \{y \in X : d(x, y) < r\}$ is an open ball. We fix a numbering $\mathscr{B} = \{B_0, B_1, \ldots\}$ which makes the number of a ball effectively computable from its center and radius and vice versa. $\mathscr{B}$ is a countable basis of the topology.

Definition 4 (Effective open sets) *We say that an open set U is* effective *if there is an algorithm $\mathscr{A} : \mathbb{N} \to \mathscr{B}$ such that $U = \bigcup_n \mathscr{A}(n)$.*

    Observe that an algorithm which diverges on each input *n* enumerates the empty set, which is then an effective open set. Sequences of uniformly effective open sets are naturally defined. Moreover, if $(U_i)_{i \in \mathbb{N}}$ is a sequence of uniformly effective open sets, then $\bigcup_i U_i$ is an effective open set.

Definition 5 (Effective G_δ-set) *An **effective G_δ-set** is an intersection of a sequence of uniformly effective open sets.*

    Obviously, an intersection of uniformly effective G_δ-sets is also an effective G_δ-set.

    Let $(X, S_X = \{s_1^X, s_2^X, \ldots\}, d_X)$ and $(Y, S_Y = \{s_1^Y, s_2^Y, \ldots\}, d_Y)$ be computable metric spaces. Let also B_i^X and B_i^Y be enumerations of the ideal balls in *X* and *Y*. A computable function $X \to Y$ is a function whose behavior can be computed by an algorithm up to any precision. For this it is sufficient that the pre-image of each ideal ball can be effectively enumerated by an algorithm.

Definition 6 (Computable Functions) *A function $T : X \to Y$ is **computable** if $T^{-1}(B_i^Y)$ is an effective open set, uniformly in i. That is, there is an algorithm $\mathscr{A} : \mathbb{N} \times \mathbb{N} \to \mathscr{B}^X$ such that $T^{-1}(B_i^Y) = \bigcup_n \mathscr{A}(i, n)$ for all i.*
*A function $T : X \to Y$ is **computable on** $D \subseteq X$ if there are uniformly effective open sets U_i such that $T^{-1}(B_i^Y) \cap D = U_i \cap D$.*

2.3 Computable measures

Let us consider the space $PM(X)$ of Borel probability measures over X. Let $C_0(X)$ be the set of real-valued bounded continuous functions on X. We recall the notion of weak convergence of measures:

Definition 7 *μ_n is said to be **weakly convergent** to μ if $\int f \, d\mu_n \to \int f \, d\mu$ for each $f \in C_0(X)$.*

Let us introduce the Wasserstein-Kantorovich distance between measures. Let μ_1 and μ_2 be two probability measures on X and consider:

$$W_1(\mu_1, \mu_2) = \sup_{f \in 1\text{-Lip}(X)} \left| \int f \, d\mu_1 - \int f \, d\mu_2 \right|,$$

where $1\text{-Lip}(X)$ is the space of functions on X having Lipschitz constant less than one.

Proposition 2 (see [1] Prop 7.1.5)

1. *W_1 is a distance and if X is bounded, separable and complete, then $PM(X)$ with this distance is a separable and complete metric space.*

2. *If X is bounded, a sequence is convergent for the W_1 metrics if and only if it is convergent for the weak topology.*

Item (1) has an effective version: $PM(X)$ inherits the computable metric structure of X. Indeed, given the set S_X of ideal points of X we can naturally define a set of ideal points $S_{PM(X)}$ in $PM(X)$ by considering finite rational convex combinations of the Dirac measures δ_s supported on ideal points $s \in S_X$. This is a dense subset of $PM(X)$. The proof of the following proposition can be found in ([11]).

Proposition 3 *If X bounded then $(PM(X), W_1, S_{PM(X)})$ is a computable metric space.*

A measure μ is then computable if there is a sequence $\mu_n \in S_{PM(X)}$ converging exponentially fast to μ in the W_1 metric (and hence μ_n weakly converge to μ).

2.4 Computable probability spaces

To obtain computability results on dynamical systems, it seems obvious that some computability conditions must be required on the system. The "good" conditions, if any, are not obvious to specify.

A computable function defined on the whole space is necessarily continuous. But a transformation or an observable need not be continuous at every point, as many interesting examples prove (piecewise-defined transformations, characteristic functions of measurable sets,...), so the requirement of being computable everywhere is too strong. In a measure-theoretical setting, a natural weaker condition is to require the function to be computable on a set of full measure. It can be proved that such a function can be extended to a function which is computable on a full-measure effective G_δ-set (see [11, 10]).

Definition 8 *A **computable probability space** is a pair (X, μ) where X is a computable metric space and μ a computable Borel probability measure on X.*

*Let Y be a computable metric space. A function $(X, \mu) \to Y$ is **almost everywhere computable** (a.e. computable for short) if it is computable on an effective G_δ-set of measure one, denoted by $\mathrm{dom}\, f$ and called the* domain of computability *of f.*

*A **morphism** of computable probability spaces $f : (X, \mu) \to (Y, \nu)$ is a morphism of probability spaces which is a.e. computable.*

Remark 1 *A sequence of functions f_n is uniformly a.e. computable if the functions are uniformly computable on their respective domains, which are uniformly effective G_δ-sets. Observe that in this case, intersecting all the domains provides an effective G_δ-set on which all f_n are computable. In the following we will apply this principle to the iterates $f_n = T^n$ of an a.e. computable function $T : X \to X$, which are uniformly a.e. computable.*

Remark 2 *The space $L^1(X, \mu)$ (resp. $L^2(X, \mu)$) can be made a computable metric space, choosing some dense set of bounded computable functions as ideal elements. We say that an integrable function $f : X \to \overline{\mathbb{R}}$ is $L^1(X, \mu)$-computable if its equivalence class is a computable element of the computable metric space $L^1(X, \mu)$. Of course, if $f = g$ μ-a.e., then f is $L^1(X, \mu)$-computable if and only if g is. Basic operations on $L^1(X, \mu)$, such as addition, multiplication by a scalar, min, max etc. are computable. Moreover, if $T : X \to X$ preserves μ and T is a.e. computable, then $f \to f \circ T$ (from L^1 to L^1) is computable (see [12]).*

2.4.1 Application to convergence of random variables

Here, (X, μ) is a computable probability space, where X is complete.

Definition 9 *A **random variable** on (X, μ) is a measurable function $f : X \to \mathbb{R}$.*

Definition 10 *Random variables f_n **effectively converge in probability** to f if for each $\varepsilon > 0$, $\mu\{x : |f_n(x) - f(x)| < \varepsilon\}$ converges effectively to 1, uniformly in ε. That is, there is a computable function $n(\varepsilon, \delta)$ such that for all $n \geq n(\varepsilon, \delta)$, $\mu\{|f_n - f| \geq \varepsilon\} < \delta$.*

Definition 11 *Random variables f_n **effectively converge almost surely** to f if $f'_n = \sup_{k \geq n} |f_k - f|$ effectively converge in probability to 0.*

The following result ([6], Theorem 2) shows that if a sequence f_n converges effectively a.s. to f then there are computable points which for which $f_n(x) \to f(x)$.

Theorem 1 *Let X be a complete metric space. Let f_n, f be uniformly a.e. computable random variables. If f_n effectively converges almost surely to f then the set $\{x : f_n(x) \to f(x)\}$ contains an effective Borel-Cantelli set (see the Appendix for the precise definition).*

In particular, it contains a sequence of uniformly computable points which is dense in $\text{Supp}(\mu)$.

Remark 3 *Moreover, the effective Borel Cantelli Set found above depends algorithmically on f_n and on the function $n(\delta, \varepsilon)$ giving the rate of convergence (see the proof of Theorem 2 in [6]). Hence the result is uniform in f_n and $n(\delta, \varepsilon)$.*

2.5 Effective L^1, L^2 convergence

Let (X, μ, T) be a computable measure-preserving system and f a L^1-computable function (in the sense that it is a computable point of the metric space L^1). It was proved in [2] that the (L^1, L^2 and almost sure) convergence of the Birkhoff averages of f is effective as soon as the norm of the limit f^* is computable. Here we give an alternative proof in the ergodic case which is simpler as it uses the classical convergence result instead of giving a "constructive" proof.

Let us call (X, μ, T) a ***computable ergodic system*** if (X, μ) is a computable probability space where T is an endomorphism (i.e. an a.e. computable measure-preserving transformation) and (X, μ, T) is ergodic. Let $||f||$ denote the L^1 norm or the L^2 norm.

Proposition 4 *Let (X, μ, T) be a computable ergodic system. Let f be a computable element of $L^1(X, \mu)$ (resp. $L^2(X, \mu)$).*

The L^1 convergence (resp. L^2 convergence) of the Birkhoff averages of f is effective.

Proof. Replacing f with $f - \int f d\mu$, we can assume that $\int f d\mu = 0$. Let $A_n = (f + f \circ T + \ldots + f \circ T^{n-1})/n$. The sequence $||A_n||$ is computable (see Remark 2) and converges to 0 by the ergodic theorems.

Given $p \in N$, we write $m \in N$ as $m = np + k$ with $0 \leq k < p$. Then

$$
\begin{aligned}
A_{np+k} &= \frac{1}{np+k} \left(\sum_{i=0}^{n-1} p A_p \circ T^{pi} + k A_k \circ T^{pn} \right) \\
||A_{np+k}|| &\leq \frac{1}{np+k} (np||A_p|| + k||A_k||) \\
&\leq ||A_p|| + \frac{||A_k||}{n} \\
&\leq ||A_p|| + \frac{||f||}{n}.
\end{aligned}
$$

Let $\varepsilon > 0$. We can compute some $p = p(\varepsilon)$ such that $||A_p|| < \varepsilon/2$. Then we can compute some $n(\varepsilon) \geq \frac{2}{\varepsilon}||f||$. The function $m(\varepsilon) := n(\varepsilon)p(\varepsilon)$ is computable and for all $m \geq m(\varepsilon)$, $||A_m|| \leq \varepsilon$. $\square$

2.6   Effective almost sure convergence

Now we use the above result to find a computable estimation for the a.s. speed of convergence.

Theorem 2 *Let (X, μ, T) be a computable ergodic system. If f is $L^1(X, \mu)$-computable, then the a.s. convergence is effective.*

This will be proved by the following

Proposition 5 *If f is $L^1(X, \mu)$-computable, and $||f||_\infty$ is bounded, then the almost-sure convergence is effective (uniformly in f and a bound on $||f||_\infty$).*

To prove this we will use the Maximal ergodic theorem:

Lemma 1 (Maximal ergodic theorem) *For $f \in L^1(X, \mu)$ and $\delta > 0$,*

$$
\mu(\{\sup_n |A_n^f| > \delta\}) \leq \frac{1}{\delta}||f||_1.
$$

The idea is simple: compute some p such that $||A_p^f||_1$ is small, apply the maximal ergodic theorem to $g := A_p^f$, and then there is n_0, that can be computed, such that A_n^f is close to A_n^g for $n \geq n_0$.

Proof. Let $\varepsilon, \delta > 0$. Compute p such that $||A_p^f|| \leq \delta\varepsilon/2$. Applying the maximal ergodic theorem to $g := A_p^f$ gives:

$$
\mu(\{\sup_n |A_n^g| > \delta/2\}) \leq \varepsilon. \tag{2.1}
$$

Now, A_n^g is not far from A_n^f: expanding A_n^g, one can check that

$$
A_n^g = A_n^f + \frac{u \circ T^n - u}{np},
$$

where $u = (p-1)f + (p-2)f \circ T + \ldots + f \circ T^{p-2}$. $||u||_\infty \leq \frac{p(p-1)}{2}||f||_\infty$ so if $n \geq n_0 \geq 4(p-1)||f||_\infty/\delta$, then $||A_n^g - A_n^f||_\infty \leq \delta/2$. As a result, if $|A_n^f(x)| > \delta$ for some $n \geq n_0$, then $|A_n^g(x)| > \delta/2$. From (2.1), we then derive

$$\mu(\{\sup_{n \geq n_0} |A_n^f| > \delta\}) \leq \varepsilon.$$

As n_0 can be computed from δ and ε, we get the result. $\square$

Remark 4 *This result applies uniformly to a uniform sequence of computable $L^\infty(X,\mu)$ observables f_n.*

We now extend this to $L^1(X,\mu)$-computable functions, using the density of $L^\infty(X,\mu)$ in $L^1(X,\mu)$.

Proof. (of Theorem 2) Let $\varepsilon, \delta > 0$. For $M \in \mathbb{N}$, let us consider $f'_M \in L^\infty(X,\mu)$ defined as

$$f'_M(x) = \begin{cases} \min(f, M) & \text{if } f(x) \geq 0 \\ \max(f, -M) & \text{if } f(x) \leq 0. \end{cases}$$

Compute M such that $||f - f'_M||_1 \leq \delta\varepsilon$. Applying Proposition 5 to f'_M gives some n_0 such that $\mu(\{\sup_{n \geq n_0} |A_n^{f'_M}| > \delta\}) < \varepsilon$. Applying Lemma 1 to $f''_M = f - f'_M$ gives $\mu(\{\sup_n |A_n^{f''_M}| > \delta\}) < \varepsilon$. As a result, $\mu(\{\sup_{n \geq n_0} |A_n^f| > 2\delta\}) < 2\varepsilon$. $\square$

Remark 5 *Also Theorem 2 applies uniformly on an uniform sequence of computable $L^1(X,\mu)$ observables f_n.*

Remark 6 *We remark that a bounded a.e. computable function, as defined in Definition 8 is a computable element of $L^1(X,\mu)$ (see [12]). Conversely, if f is a computable element of $L^1(X,\mu)$ then there is a sequence of uniformly computable functions f_n that effectively converge μ-a.e. to f.*

3 Pseudorandom points and dynamical systems

As said before the famous Birkhoff ergodic theorem says that in an ergodic system, the time average computed along μ-almost every orbit coincides with space average with respect to μ.

If a point x satisfies Equation 1.1 for a certain f, then we say that x is typical with respect to the observable f.

Definition 12 *A point x is μ-**typical** if x is typical w.r.t. every continuous function $f : X \to \mathbb{R}$ with compact support.*

We remark that *from now on we will suppose that X is a complete metric space.* We will see that such μ−typical points exist in computable ergodic systems. First we give a result for L^1 observables.

Theorem 3 *If (X,μ,T) is a computable ergodic system, f is $L^1(X,\mu)$ and a.e. computable then there is a uniform sequence x_i of computable points which is dense on the support of μ such that for each i*

$$\lim_{n \to \infty} \frac{1}{n} \sum f(T^n(x_i)) = \int f \, d\mu.$$

Proof. Apply theorem 1 to the sequence of uniformly a.e. computable functions $f_n = A_n^f$ which converge effectively almost-surely by theorem 2. We obtain that the set of points for which $\frac{1}{n}\sum f(T^n(x_n)) \to \int f \, d\mu$ contains a sequence of computable points, as in the statement. $\square$

Let g be the point-wise limit of a sequence of uniformly computable functions f_n (defined μ-a.e., as in Remark 5): Theorem 3 can also be proved to hold for the observable g, i.e. in a computable ergodic system there exists computable points for which the Birkhoff averages of g converge to $\int g \, d\mu$.

Since it is possible to construct a r.e. set of computable functions which is dense in the space of compactly supported continuous functions we can also obtain the following

Theorem 4 *If (X, μ, T) is a computable ergodic system then there is a uniform sequence x_n of computable points which in dense on the support of μ such that for each n, x_n is μ-typical.*

Proof. Let us introduce (following [7]) a certain fixed, enumerated sequence of Lipschitz functions. Let $\mathscr{F}_0$ be the set of functions of the form:

$$g_{s,r,\varepsilon} = |1 - |d(x,s) - r|^+/\varepsilon|^+ \tag{3.1}$$

where $s \in S$, $r, \varepsilon \in \mathbb{Q}$ and $|a|^+ = \max\{a, 0\}$.

$g_{s,r,\varepsilon}$ is a Lipschitz functions whose value is 1 inside the ideal ball $B(s,r)$, 0 outside $B(s, r + \varepsilon)$ and with intermediate values in between. It is easy to see that the real-valued functions $g_{s_i, r_j, \varepsilon_k} : X \to \mathbb{R}$ are computable, uniformly in i, j, k.

Let $\mathscr{F}$ be the smallest set of functions containing $\mathscr{F}_0$ and the constant 1, and closed under max, min and rational linear combinations. Clearly, this is also a uniform family of computable functions. We fix some enumeration $\nu_{\mathscr{F}}$ of $\mathscr{F}$ and we write g_n for $\nu_{\mathscr{F}}(n) \in \mathscr{F}$. We remark that this set is dense in the set of continuous functions with compact support.

By Remark 6 moreover these are computable elements of $L^1(X, \mu)$, hence Theorem 2 applies uniformly to these observables. This means that we can apply Theorem 1 uniformly on this sequence. By intersecting all effective Borel-Cantelli sets given by Theorem 1, since the intersection of a uniform family of effective BC sets contains an effective BC set (see Remark 3 and Proposition 6) and such a set contains a sequence of computable points which are dense in the support of μ (see Theorem 5), which are typical *for all* observables in $\mathscr{F}$ (in the same way as in Theorem 3). Since each continuous function f with compact support can be approximated in the L^∞ norm by a function in $\mathscr{F}$ (and in particular the approximating function in $\mathscr{F}$ have values near the values of f at almost each point) the statement is proved. □

3.1   Conclusion and some open question

We have seen that in computable ergodic systems, the speed of a.e. convergence of ergodic averages is computable, and if moreover also the invariant measure is computable then there are computable points which are typical for the statistical behavior of the system: the pseudorandom points.

The assumption about the computability of the measure is not redundant with the assumption about the computability of the system because, as said before, there are computable systems having not computable invariant measures. It is also interesting to remark that there are systems where the map T is computable, for which we can consider a non computable ergodic measure having no pseudorandom points (an example can be constructed considering a Bernoulli shift where the symbols have non computable probability).

Our results about the existence of these points, however, tell not much about the computational complexity which is necessary to find them. It would be of practical importance to have fast algoritms for this computation.

All these questions are related to another general (and vague) question we like to cite, which is of great practical importance: why, many "naive" simulations of dynamical systems give reasonable results? or more precisely: under which assumptions the simulations of a dynamical system by a computer (which is a kind of discrete model for a continuous phenomena) give correct results?

4   Appendix: effective BC sets

We recall some results from [6] which are used in the proofs of the present paper.

Given a measurable space X endowed with a probability measure μ, the well known Borel-Cantelli lemma states that if a sequence of sets A_k is such that $\sum \mu(A_k) < \infty$ then the set of points which belong to finitely many A_k's has full measure. It holds that if the A_k are given in some "effective" way (and μ is computable) then this full measure set contains computable points, which can be effectively constructed.

Definition 13 *A sequence of positive numbers a_i is* effectively summable *if the sequence of partial sums converges effectively: there is an algorithm $\mathscr{A} : \mathbb{Q} \to \mathbb{N}$ such that if $A(\varepsilon) = n$ then $\sum_{i \geq n} a_i \leq \varepsilon$.*

For the sake of simplicity, we will focus on the complements U_n of the A_n.

Definition 14 *An* effective Borel-Cantelli sequence *is a sequence $(U_n)_{n \in \mathbb{N}}$ of uniformly effective open sets such that the sequence $\mu(X \setminus U_n)$ is effectively summable.*
The corresponding effective Borel-Cantelli set *is $\bigcup_k \bigcap_{n \geq k} U_n$.*

Proposition 6 *The intersection of any uniform family of effective Borel-Cantelli sets contains an effective Borel-Cantelli set.*

Theorem 5 *Let X be a complete Computable Metric Space and μ a computable Borel probability measure on X.*
Every effective Borel-Cantelli set R, contains a sequence of uniformly computable points which is dense in the support of μ.

References

[1] L. Ambrosio, N. Gigli, G. Savare. *Gradient Flows: In Metric Spaces And In She Space Of Probability Measures*, Birkhauser Zurich 2005 ebnisse der Mathematik und ihrer Grenzgebiete.

[2] J. Avigad, P. Gerhardy, H. Towsner. *Local stability of ergodic averages.* Transactions of the American Mathematical Society, **362** (2010) 261–288

[3] V. Brattka , P. Hertling, K. Weihrauch *A Tutorial on Computable Analysis* in New Computational Paradigms, Springer New York 2007

[4] P. Gács. *Lectures notes on descriptional complexity and randomness.* Boston University (1993) 1–67.

[5] P. Gács, M. Hoyrup, C. Rojas. *Randomness on computable probability spaces - a dynamical point of view.* In Susanne Albers and Jean-Yves Marion, editors, *26th International Symposium on Theoretical Aspects of Computer Science (STACS 2009)*, 469–480

[6] S. Galatolo, M. Hoyrup, and C. Rojas. *A constructive Borel-Cantelli lemma. Constructing orbits with required statistical properties.* Theor. Comput. Sci., **410**, (2009) 2207–2222,.

[7] Peter Gács. Uniform test of algorithmic randomness over a general space. *Theoretical Computer Science*, 341:91–137, 2005.

[8] Stefano Galatolo, Mathieu Hoyrup, and Cristóbal Rojas. *Dynamics and abstract computability: computing invariant measures.* arXiv:0903.2385 . To appear on Disc. Cont. Dyn. Sys. A. Available at `http://front.math.ucdavis.edu/0903.2385`

[9] Boris Hasselblatt and Anatole Katok. *Introduction to the Modern Theory of Dynamical Systems*, volume 54 of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, 1995.

[10] Mathieu Hoyrup. *Computability, randomness and ergodic theory on metric spaces.* PhD thesis, Université Paris Diderot (Paris 7), 2008.

[11] M. Hoyrup and C. Rojas. *Computability of probability measures and Martin-Löf randomness over metric spaces.* Inf. Comput., (2009) **207**, 830–847.

[12] Hoyrup, M., Rojas, C.: An application of Martin-Löf randomness to effective probability theory. In: LNCS. Proceedings of CiE'09.

[13] H. Rogers. *Theory of recursive functions and effective computability* MIT Press Cambridge, MA, USA (1987)

[14] Cristóbal Rojas. *Randomness and ergodic theory: an algorithmic point of view*. PhD thesis, Ecole Polytechnique, 2008.

[15] Alan Turing. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, 2, 42:230–265, 1936.

[16] K. Weihrauch. *Computable analysis. An introduction.*, Springer, 2000.

[17] Lai-Sang Young. What are SRB measures, and which dynamical systems have them? *J. Stat. Phys.*, 108:733–754, 2002.