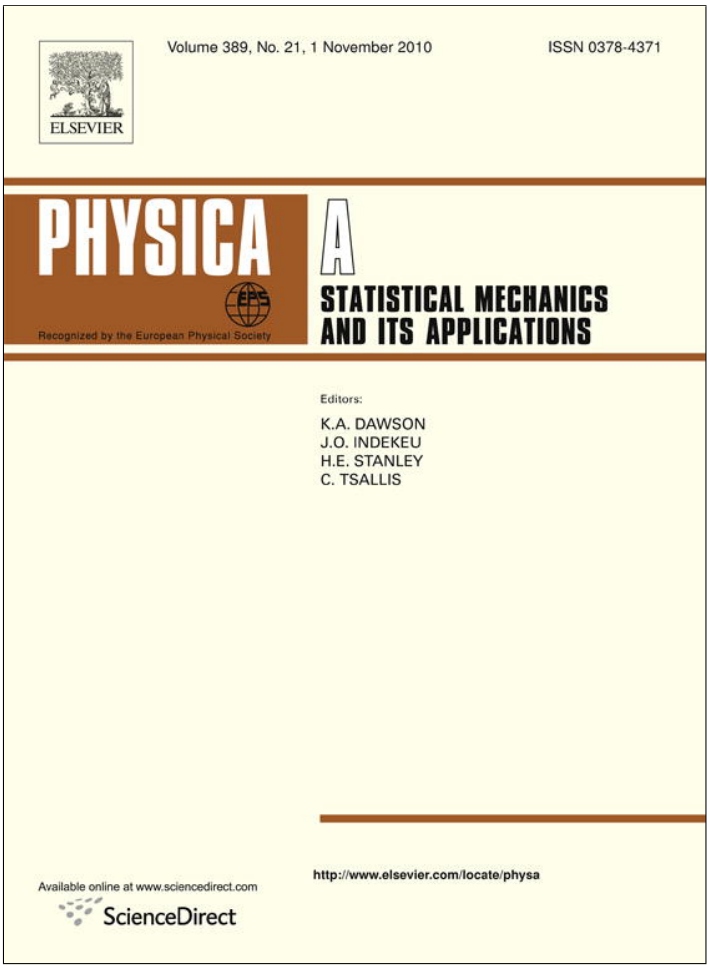

Physica A 389 (2010) 4994–5006

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Asymptotic distribution of global errors in the numerical computations of dynamical systems

G. Turchetti [a], S. Vaienti [b], F. Zanlungo [c,d,*]

[a] Department of Physics, University of Bologna and INFN, Bologna, Italy
[b] Centre de Physique Théorique, CNRS, Universités d'Aix-Marseille I, II, Université du Sud, Toulon-Var and FRUMAM, France
[c] ATR Intelligent Robotics and Communication Laboratories, Kyoto, Japan
[d] Department of Physics, University of Bologna, Italy

ARTICLE INFO

Article history:
Received 5 April 2010
Received in revised form 29 June 2010
Available online 23 July 2010

Keywords:
Discrete systems
Effect of numerical round-off

ABSTRACT

We propose an analysis of the effects introduced by finite-accuracy and round-off arithmetic on numerical computations of discrete dynamical systems. Our method, which uses the statistical tool of the decay of fidelity, computes the error by directly comparing the numerical orbit with the exact one (or, more precisely, with another numerical orbit computed with a much higher accuracy). Furthermore, as a model of the effects of round-off arithmetic on the map, we also consider a random perturbation of the exact orbit with an additive noise, for which exact results can be obtained for some prototype maps. We investigate the decay laws of fidelity and their relationship with the error probability distribution for regular and chaotic maps, for both additive and numerical noise. In particular, for regular maps we find an exponential decay for additive noise, and a power-law decay for numerical noise. For chaotic maps, numerical noise is equivalent to additive noise, and our method is suitable for identifying a threshold for the reliability of numerical results, i.e., the number of iterations below which global errors can be ignored. This threshold grows linearly with the number of bits used to represent real numbers.

1. Introduction

The reliability of numerical computations in dynamical systems is a relevant long-standing question. For continuous systems, the global error on the orbit is the result of the accumulation of two types of error: the first is the local error due to the discretization algorithms, the second is due to the finite-precision representation of reals in any computational device and to the round-off arithmetic. In this paper, we propose a statistical analysis of the global errors due to finite accuracy and round-off in discrete dynamical systems. Even though our results on numerical maps are based on numerical computations, the comparison with randomly perturbed maps, for which analytical results are available, allows to understand the key features. For maps with hyperbolic attractors, the existence of true orbits close to pseudo-orbits as established by the shadowing lemma [1–8] is of great theoretical importance but does not provide any statistical information on the discrepancies between a pseudo-orbit and a true orbit with the same initial condition. The perturbations of dynamical systems have also been investigated in the context of algorithmic complexity in [9–11]. We show that for chaotic maps there is a sharp transition between two regions, one in which the error is negligible and one in which its variance is comparable with the size of the attractor, for both the numerical and the randomly perturbed map. For regular or quasi-regular anisochronous maps, the transition is smooth, and the accuracy decay follows a power law for the numerical map and exponential law for the randomly perturbed map.

* Corresponding author at: ATR Intelligent Robotics and Communication Laboratories, Kyoto, Japan.
E-mail address: francesco.zanlungo@gmail.com (F. Zanlungo).

0378-4371/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2010.06.060

In a previous work [12], we compared the value of an observable computed along the orbit of a given map and the orbit of the same map with a random perturbation for the same initial condition. The difference is a random process and, for large iteration times, its probability density function (pdf) converges toward a widespread distribution, which depends on the choice of the observable. For chaotic maps, the relaxation is sharp, in the sense that there exists a threshold time below which the width of the pdf is very small, approximating a δ function, and above which it has already reached its asymptotic value. These results were obtained analytically and numerically by computing the inverse Fourier transform of a suitable correlation integral between the two observables, called *fidelity*. For chaotic systems, the fidelity decays at least exponentially quickly, and this allows one to recover the average of the observables with respect to the invariant and the stationary measures. For Bernoulli maps, the decay is super-exponential, and there is numerical evidence for this type of decay for most chaotic maps numerically investigated. For regular maps such as the translations on the torus, the decay is smoother, and it follows an exponential law. Numerical evidence was found that the same exponential decay persists for quasi-integrable maps.

In a subsequent letter [13], we applied the previous theoretical framework to compare the exact orbit of a map with its numerical computation (on a machine). The latter could be considered as a pseudo-random orbit where the noise is given by the round-off errors.

In the present paper, we develop with more details and formal considerations the results announced in [13], and provide some new results.

The difference between the values of an observable computed along the exact and the numerical orbits shows again the existence of an abrupt transition time n^* for the corresponding pdf, which has an interpretation of practical interest. We could in fact argue that below the transition time the numerical orbits are *faithful* to the exact orbit. By the way, this approach requires the knowledge of the exact map, which is not accessible. This difficulty is overcome by comparing the numerical map computed for a given accuracy (single floating-point precision) with the same map evaluated numerically with a higher accuracy (double or higher floating-point precision), that we consider as the *reference* map. We show that the results of the single–higher-precision comparison are equivalent, at least below a significant time scale, to the results we would obtain by comparing a single-precision map and the exact map. We show that the threshold n^* grows like $-\ln\varepsilon$, where ε is the accuracy specified by the least significant bit used to represent a real number. As a consequence, the fidelity threshold n^* grows linearly with the length of the string of bits used to represent real numbers. Beyond this threshold, the error distribution spreads quickly over its accessible range.

The behavior just described is typical of chaotic systems. For regular maps the situation is slightly different: for isochronous maps the fidelity error does not decay, whereas for anisochronous maps the decay follows a power law. For these systems it is not possible to identify a sharp transition from the delta function to the asymptotic one; the transition is a gradual process whose length depends on ε, i.e., it gets longer with the number of bits used to represent real numbers. A similar decay occurs for the correlations where it is due to the local mixing properties, a property typical of anisochronous maps.

The plan of the paper is the following. In Section 2, we review some results on additive noise obtained by the authors of this paper in a previous work. In Section 3, we introduce the problem of numerical noise and the notation we are using in order to handle it. In Section 4, fidelity is generalized in order to study numerical noise. We study the effect of numerical noise on regular maps in Section 5 and on chaotic maps in Section 6. Our results are summarized in the conclusions.

2. Additive noise

We recall in this section the main results obtained applying the fidelity to random perturbations of dynamical systems, results that have been presented in [12]. We consider a map T defined on a phase space X which is a subset of $\mathbb{R}^d$, endowed with an invariant *physical* measure μ defined by

$$\lim_{n\to\infty}\int_X \Phi(T^n(x))\mathrm{d}m(x) = \int_X \Phi(x)\mathrm{d}\mu(x), \tag{1}$$

where m denotes the Lebesgue measure and Φ is a continuous observable. Let us then consider a sequence of independent and identically distributed random variables ξ_i with values in the probability space Ξ and with probability density distribution $\eta(\xi)$ such that $Tx+\varepsilon\xi$ still maps X into itself. The iteration of the map T is therefore replaced by a composition of maps chosen randomly close to it (note that T itself is recovered when $\varepsilon = 0$): $T_\varepsilon^n(x) = (T+\varepsilon\xi_n)\circ(T+\varepsilon\xi_{n-1})\circ\cdots\circ(T+\varepsilon\xi_1)(x)$ and the stationary measure μ_ε of the process is defined by [14]:

$$\lim_{n\to\infty}\int_{X,\Xi}\mathrm{d}m(x)\prod_i \eta(\xi_i)\mathrm{d}\xi_i\Phi(T_\varepsilon^n(x)) = \int_X \Phi(x)\mathrm{d}\mu_\varepsilon(x). \tag{2}$$

We want to study the statistical properties of the error at the n-th iteration, defined as $\Delta_\varepsilon^n(x) = f(T^n x) - f(T_\varepsilon^n x)$, where f is a smooth observable; in the following, we take $f(x) = x$ for one-dimensional maps and $f(\mathbf{x}) = x_i,\ i = 1,\ldots,d, \mathbf{x}\equiv(x_1,\ldots,x_d)$, for multi-dimensional maps. The error Δ_ε^n can be considered as a random variable on the probability space given by the direct product of X (space of the initial conditions) and Ξ^∞ (space of the realizations), the former endowed with the Lebesgue measure m, and the latter with the product measure $\prod_i \eta(\xi_i)\mathrm{d}\xi_i$. We are in particular interested in studying the

probability distribution function (pdf) ρ_ε^n of the random variable $\Delta_\varepsilon^n(x)$, which could be studied directly, through a Monte Carlo sampling over initial conditions x and random perturbations ξ, or indirectly, using the fidelity. In fact, the fidelity integral will allow us to compute such a pdf as the Fourier inverse transform of the characteristic function of Δ_ε^n.

Fidelity is defined through the following integral:

$$F_\varepsilon^n = \int_{X,\Xi} \mathrm{d}m(x) \prod_i \eta(\xi_i)\mathrm{d}\xi_i \Phi(T^n(x))\Psi(T_\varepsilon^n(x)), \tag{3}$$

where also Ψ, as Φ, is an arbitrary continuous function.

For a large class of maps which mix exponentially quickly, it can be shown that the fidelity converges to $\int_X \Phi(x)\,\mathrm{d}\mu(x) \int_X \Psi(x)\mathrm{d}\mu_\varepsilon(x)$, and we define as the *fidelity error*, denoted δF_ε^n, the absolute error of the difference between the integral (3) and its limit value. Note that by the asymptotic characterization of the invariant and stationary measures, the fidelity error could be equivalently defined as

$$\delta F_\varepsilon^n = F_\varepsilon^n - \int_X \Phi(T^n(x))\mathrm{d}m(x) \int_{X,\Xi} \prod_i \eta(\xi_i)\mathrm{d}\xi_i \mathrm{d}m(x)\Psi(T_\varepsilon^n(x)), \tag{4}$$

which emphasizes the role of the Lebesgue measure in the numerical computations. This quantity is relevant because the inverse Fourier transform of the fidelity $F_\varepsilon^n(u)$, computed by choosing $\Phi(x) = \mathrm{e}^{\mathrm{i}uf(x)}$ and $\Psi(x) = \mathrm{e}^{-\mathrm{i}uf(x)}$, is the probability error distribution ρ_ε^n. By definition, when the fidelity error converges to zero, then the fidelity converges to the product of the Fourier transforms of the measures $\int_X \mathrm{e}^{\mathrm{i}ux}\,\mathrm{d}\mu(x)\ \int_X \mathrm{e}^{-\mathrm{i}ux}\,\mathrm{d}\mu_\varepsilon(x)$, and its inverse Fourier transform is just ρ_ε^∞. The initial error distribution is the Dirac distribution $\rho^0(s) = \delta(s)$, whereas it can be shown that if the invariant and stationary measures are Lebesgue the asymptotic distribution is the triangular function $\rho_\infty(s) = (1 - |s|)\vartheta(1 - |s|)$. Since the decay time scales of the fidelity error depend very weakly on the choice of the observables Φ and Ψ, the study of δF_ε^n using $\Phi(x) = \Psi(x) = x$ is usually the most efficient tool to investigate the convergence to the asymptotic error distribution ρ_ε^∞ (these observables have been used to obtain all the figures in this work, if not specified differently).

For a couple of systems, translations on the torus (a regular system) and the Bernoulli map (a chaotic system), analytical results are available also for the transient. Choosing the probability distribution $\eta(\xi) = \frac{1}{2}\chi_{[-1,1]}(\xi)$ we found that the fidelity for translations is given by

$$F_\varepsilon^n = \sum_{k\in\mathbb{Z}} \Phi_k \Psi_{-k} S^n(k\varepsilon) \qquad S(x) = \frac{\sin(2\pi x)}{2\pi x}, \tag{5}$$

where Φ_k and Ψ_k are the Fourier components of functions Φ and Ψ, while for the Bernoulli map $Tx = qx \bmod 1$, with integer $q \geq 2$ we have

$$F_\varepsilon^n = \sum_{k\in\mathbb{Z}} \Phi_k \Psi_{-k} S_{n,q}(k\varepsilon) \qquad S_{n,q}(x) = \prod_{j=0}^{n-1} S(q^j x). \tag{6}$$

In the first case, the decay of fidelity is exponential, with time scale ε^{-2}, while in the second case we have a plateau of length $n_* \propto -\ln\varepsilon$, below which the fidelity is almost constant, followed by an ε-independent super-exponential decay. Below the threshold n_*, the error probability distribution can be approximated by a δ function, and the perturbed system can be considered as equivalent to the unperturbed one. The asymptotic error distribution is the same for the two systems (since the physical and stationary measures coincide with Lebesgue), and results in being the triangular function (Figs. 1 and 2). The transition between the δ and the triangular functions is sharp and independent of ε for the Bernuolli system, while it is gradual and dependent on ε for the translations (see [12]).

Our numerical study of maps for which analytical results are not available (Hénon, Baker, Intermittent, Logistic and Standard maps) [12] shows that the behavior of translations and of Bernoulli maps can be considered as a prototype of, respectively, regular and chaotic maps. In particular, for all the studied chaotic maps it is possible to identify a threshold $n_* \propto -\ln\varepsilon$ below which the perturbed system can be considered as faithful to the unperturbed one: a threshold followed by an ε-independent, super-exponential decay. To this threshold corresponds a sharp transition of the error probability distribution from a δ function to the asymptotic distribution (that in general is not a triangular function but depends on the invariant and stationary measures).

3. Numerical noise

In this section, we review the basics of numerical round-off and introduce the notation that we are going to use to handle it in similarity with our previous study on additive noise.

Since real numbers have to be represented as strings of bits on a computer, to each discrete map T there corresponds a numerical map T_*. The action of the numerical map depends on the length of bit strings used to represent real numbers and on the details of round-off algebra, which are hardware dependent [15] (even if nowadays the differences between different

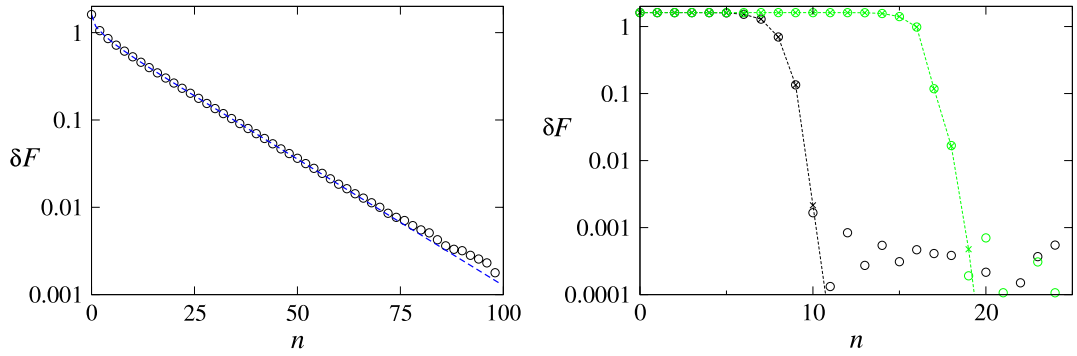

Fig. 1. Comparison between analytical results and Monte Carlo integrals for the decay of fidelity. Left: translations, $\varepsilon = 0.1$ (analytical result: blue dashed line; Monte Carlo: circles). Right: $3x$ mod 1. $\varepsilon = 10^{-4}$ in black and $\varepsilon = 10^{-8}$ in green (analytical result: dashed line and crosses; Monte Carlo: circles). We used $N = 10^7$ integration points in the Monte Carlo method. The test functions $\Phi = \Psi^*$ were defined in Fourier space as $\Phi_k = k^{-1}$, truncated at $k = 30$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

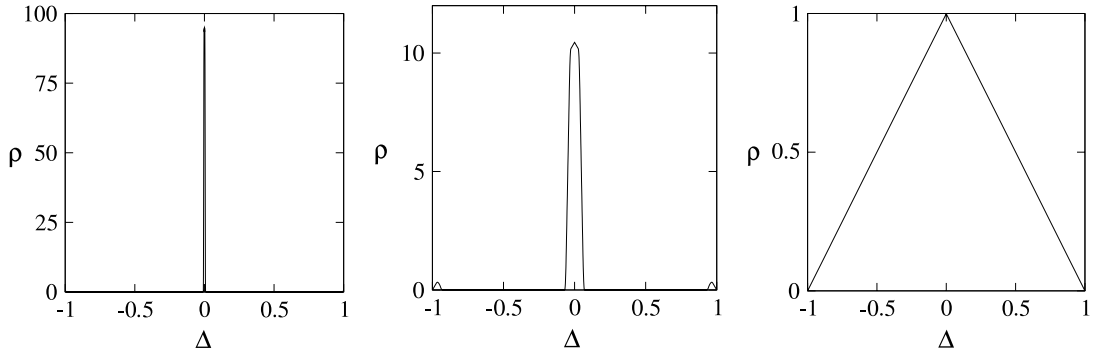

Fig. 2. ρ_ε^n for $3x$ mod 1, $\varepsilon = 10^{-8}$. Left: $n = 13$; center: $n = 15$, right: $n = 19$. Compare the transition times with Fig. 1, right.

architectures are limited by the adoption of the IEEE-754 rounding standard); nevertheless, some general results can be stated. In our computational device, any real number $x \in \mathbb{R}$ is represented by a floating-point number x^{*p}, that is, a binary string with p significant bits:

$$x^{*p} = f2^e \qquad f = \pm 0.f_1 f_2 \cdots f_p = \sum_{k=1}^{p} f_k 2^{-k} \qquad e = \pm \sum_{k=1}^{q} e_k 2^k,$$

where f_k, e_k are 0 or 1. The suffix p will be neglected whenever there is no ambiguity. The result of any arithmetic operation such as $x^{*p} \oplus y^{*p}$ corresponding to $x + y$ is a floating-point number $z^{*p} \in \mathbb{F}_p$, and sum $\oplus$ implies a round-off. The relative error r_p is defined by

$$x^{*p} = x(1 + r_p) \qquad |r_p| \le \epsilon = 2^{-p}.$$

For instance, in floating-point single precision we have $\epsilon = 2^{-25}$. We remark that the previous equation shows that, since the relative error is bound, the absolute error strongly depends on the value of the number x. This is a significant difference from the case of additive noise, which causes the floating-point discretization to be non-uniform (as we will show later, typical error distributions due to numerical round-off are stepwise). Using the notation introduced for additive noise, we define by $T_* \equiv T_\epsilon$ the map with round-off, so that letting $x_1^* = T_*(x)$ and $x_1 = T(x)$ be the images of a point x with and without round-off of amplitude ϵ, the error is

$$\epsilon\xi_1 = x_1^* - x_1 = T_*(x) - T(x),$$

where $|\xi_1| \le 1$. Unlikely the additive noise ξ_1 depends on x, so that we can write

$$T_\epsilon(x) \equiv T_*(x) = T(x) + \epsilon\xi(x).$$

As a consequence, the error at the first iteration is $\epsilon\xi_1 = \epsilon\xi(x)$. After n iterations, the local error is given by

$$\epsilon\xi_n = x_n^* - T(x_{n-1}^*) = T_*(x_{n-1}^*) - T(x_{n-1}^*), \tag{7}$$

and from the previous relation we have

$$\xi_n = \xi(x_{n-1}^*).$$

The global error is the cumulative result of the local errors, and it is defined according to

$$\Delta_n^* = x_n^* - x_n = T_*^n(x) - T^n(x).$$

We can estimate it by using the definition of the local error given by (7). Indeed,

$$x_n^* = (T+\epsilon\xi)\circ(T+\epsilon\xi)\circ\cdots(T+\epsilon\xi) = x_n + \epsilon\xi(x_{n-1}) + \epsilon\sum_{k=1}^{n-1} DT^{n-k}(x_k)\xi(x_{k-1}) + O(\epsilon^2),$$

where $DT(x)$ is the Jacobian matrix of the map T. Observing that $x_k^* = x_k + O(\epsilon^2)$, we have $\xi(x_k) = \xi(x_k^*) + O(\epsilon^2) = \xi_k + O(\epsilon^2)$, so we can write

$$\Delta_n^* = \epsilon\xi_n + \epsilon\sum_{k=1}^{n-1} DT^{n-k}(x_k)\xi_k + O(\epsilon^2). \tag{8}$$

The same expression is obtained in the case of additive noise, where

$$x_n^* = (T+\epsilon\xi_1)\circ(T+\epsilon\xi_2)\circ\cdots(T+\epsilon\xi_n).$$

However, for the additive noise, the ξ_k are independent random variables, whereas for the round-off the ξ_k depend on the initial condition according to $\xi_k = \xi(x_{k-1}^*) = \xi(x_k) + O(\epsilon^2)$. For the translations on the torus $DT = 1$, and we can write (8) as

$$\Delta_\varepsilon^n = \epsilon\sum_{k=1}^{n}\xi_k = \epsilon(n\overline{\xi} + w_n), \tag{9}$$

where $\overline{\xi} \equiv \langle\xi\rangle$ is the average of the stochastic process and w_n are the fluctuations. If $\overline{\xi} \gg w_n$ (we will see that they are of order 1 for round-off noise) we have $\lim_{n\to\infty}\Delta_\varepsilon^n/n = \epsilon\overline{\xi}$. The random perturbations we consider usually have zero mean, so $\overline{\xi} = 0$; this may not be the case for round-off errors, whose average can be non-zero. In the case of Bernoulli maps, DT is a constant, and if the process ξ has zero mean again $\overline{\xi} = 0$, unlike the case of round-off errors discussed below. We also outline that, when dealing with round-off errors, the summation in Eq. (8) starts with $k = 0$ because the initial error $x_0 - x_0^*$ is also present. Here x_0^* stands for the finite-accuracy representation of the initial condition x_0, which is in general different from x_0. As we will see, this may have relevant consequences.

4. Generalizations of fidelity

In this section we introduce a couple of new observables, in Eqs. (12) and (13) below, that we also call fidelity. To mark their difference from the observable defined above in Eq. (3), notice that this latter is defined by an integral over all the possible single-step error realizations ξ_i. Nevertheless, from a numerical point of view, integrals were performed with a Monte Carlo method, i.e., choosing N representative random vectors (x, ξ_i), a procedure that led to a relative error of order $N^{-1/2}$ (see the discrepancy between the numerical and analytical results in Fig. 1, in particular for low values of δF).

We suppose that if the deterministic $\xi_i(x)$ functional dependence of single-step errors on the initial condition is complex enough, the vector $(x, \xi_i(x))$ can be considered as equivalent to a random sequence, i.e. the Monte Carlo integral over initial conditions and noise can be substituted for round-off by an integral over the only initial conditions.

Corresponding to this *ansatz*, whose validity we are going to verify, we can compute the fidelity error for a system perturbed with numerical noise as

$$F_*^n = \int_X \Phi(T^n x)\Psi(T_*^n x)\mathrm{d}m(x) \tag{10}$$

$$\delta F_*^n = F_*^n - \left(\int_X \Phi(T^n x)\mathrm{d}m(x)\right)\left(\int_X \Psi(T_*^n x)\mathrm{d}m(x)\right). \tag{11}$$

This definition requires the knowledge of the exact map T, which is in general not available. This problem can be solved by comparing the round-off map T_*, realized with a given precision, i.e. as a string of bits of a given length, with a map realized at an higher precision, $T_\dagger$, that we call the "reference" map. For example T_* could be a single precision (8 digits) map, and $T_\dagger$ a double-precision (16 digits) map. The numerically computed fidelity will thus be

$$F_*^n = \int_X \Phi(T_\dagger^n x)\Psi(T_*^n x)\mathrm{d}m(x). \tag{12}$$

To check the relevance of these results we can compare them with those obtained substituting $T_\dagger$ with $T_\ddagger$ (for example a map realized using 24 or 32 significant digits). If the results do not depend on the precision of the reference map, we can assume that they are equivalent to those that could be obtained if we had access to the exact map. We have applied this procedure to the results shown in this paper. Typically, the results obtained using as reference map a double-precision $T_\dagger$ or a 24 (32) digit map $T_\ddagger$ are equivalent below a given time scale. This time scale, that corresponds to the time scale under which $T_\dagger$ can be considered equivalent to $T_\ddagger$, as can be checked through a direct comparison between $T_\dagger$ and $T_\ddagger$, is considerably longer than the time scale at which the error probability distribution of T_* has reached its asymptotic form and thus the results

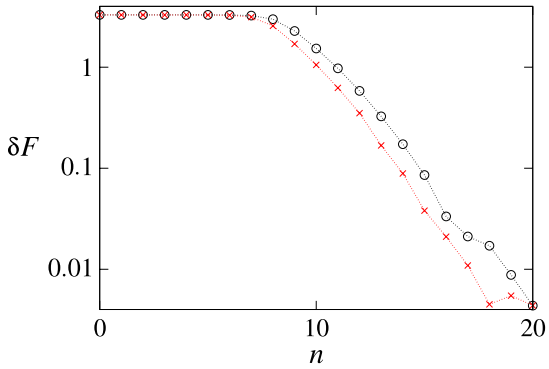

Fig. 3. Decay of fidelity for the standard map with $K = 10$, $\varepsilon = 10^{-8}$, Monte Carlo integrals using $N = 10^6$; red and crosses: comparison between single and double precision, black and circles: single precision using inversion.

of the single–double precision comparison can be considered equivalent to those that would be obtained by a comparison between a single-precision map and an exact map.

Another way to avoid the problems related to the inaccessibility of the exact map would be to rely, for invertible maps, on a different definition of fidelity as

$$\tilde{F}_*^n = \int_X \Phi(x)\Psi(T_*^n T_*^{-n} x)\mathrm{d}m(x) \tag{13}$$

where T_*^{-1} is the numerical realization of the inverse map, which is in general different from the inverse of T_*. Note that in the deterministic setting and with invertible maps T_* preserving the measure m, Eq. (13) is equivalent to (3) [16]. We generalize it by integrating on given different realizations respectively of T_*^{-n} and T_*^n, and by supposing that this integral converges to $\int \Phi \mathrm{d}m \int \Psi \mathrm{d}m$ with the same rate as the fidelity error (11). For the numerical realization of an invertible map as the standard map, the equivalence between the two definitions has been checked with good results, at least in the chaotic regime (Fig. 3).

We have performed our study using three different architectures: a processor 2 GHz Intel Core Duo on a Mac OSX v10.5 (Darwin 9) operating system, a processor Intel Pentium 4 3.00 GHz on a Linux Fedora 4 OS, and a processor 1 500 MHz IP35 on a IRIX 6.5 OS (when not specified, the results shown in figures are obtained on the Darwin architecture). The details of the round-off process depend on the architecture, nevertheless our studies show that some general rules about the error distribution can be stated.

5. Regular maps

We now proceed to the analysis of regular maps.

For regular maps the behavior is significantly different from additive noise, showing that the integral over the only initial conditions is not equivalent for these systems to an integral also on the noise. For translations on the torus, $Tx = x+\omega \bmod 1$, which are the prototypes for integrable maps, we have found that is possible to write the global error $\Delta_*^n(x,\omega) = T^n x - T_*^n x$ as in (9)

$$\Delta_*^n(x,\omega) = \varepsilon(\overline{\xi}(\omega)\phi(n) + w_n(x)), \quad \phi(n) = n + \phi_0 + \phi_1 n^{-1} + \cdots. \tag{14}$$

Here ε is a constant that represents the least significant bit accuracy, $\overline{\xi}$ is a constant that depends in a non-trivial (and machine-dependent) way on ω, which as we will see is equivalent to an average of the single-step errors ξ_i, while w_n is a bounded, periodic function which depends on the initial condition x and has zero average with respect to the initial conditions. Notice that in this case $\overline{\xi}$ is not zero, in contrast with the case of random noise. We thus have

$$\lim_{n\to\infty} \frac{\Delta_*^n(x,\omega)}{n} = \varepsilon\overline{\xi}(\omega). \tag{15}$$

Fig. 4 shows the global error $\Delta_*^n(x,\omega)$ for two different initial conditions x and the same ω, on two different time scales, showing that the contributions of w_n go to zero in the limit. The periodic nature of the single-step errors ξ_i (and thus of w_n) is shown by Fig. 5 (periodicity emerges after some non-periodic iterations). The "asymptotic average error" $\overline{\xi}(\omega)$ is clearly correlated to the round-off error $\delta\omega = \omega - \omega_*$ of the translation parameter, as shown by Fig. 6. This figure shows at the left the distribution of $\overline{\xi}(\delta\omega)$ and at the right the $\overline{\xi}$ and $\delta\omega$ distributions obtained using a Monte Carlo sampling of the parameter ω. From the figure at the left we can see that $\overline{\xi}(\delta\omega)$ is given by a distribution law centred around $\overline{\xi} = \delta\omega$. The distribution law for the round-off error on the translation parameter

$$\eta(\delta\omega) = 2^{p-1}\sum_{k=p}^{\infty} \chi_{[-2^{-k},2^{-k}]}(\delta\omega),$$

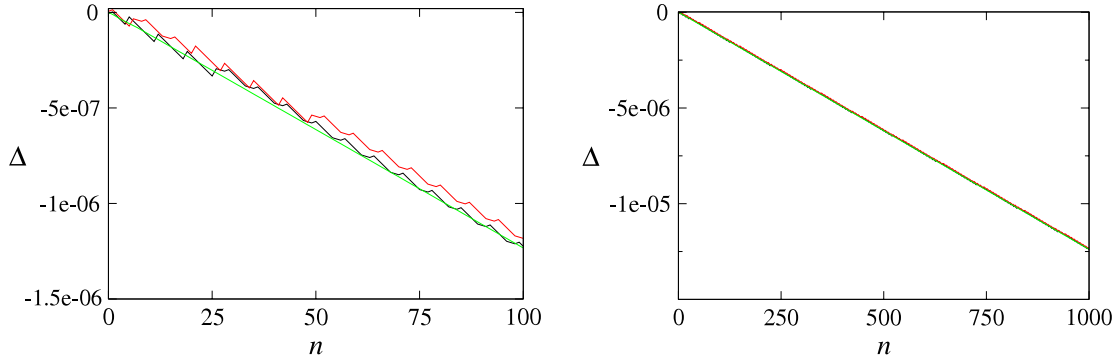

Fig. 4. $\Delta_*^n(x, \omega)$ in translations on the torus for two different, randomly chosen, initial conditions x (red and black), compared with an average over 10^4 values of x (green). $\omega = \sqrt{2}/10$. The two figures show the same results on different time scales. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

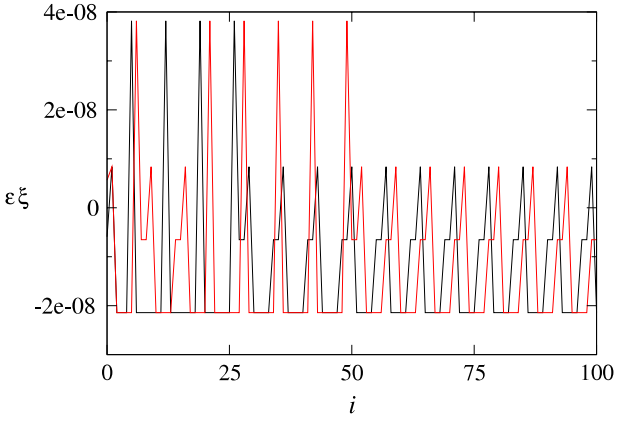

Fig. 5. Single-step errors ξ_i, for two different, randomly chosen, initial conditions x. Translations on the torus, $\omega = \sqrt{2}/10$.

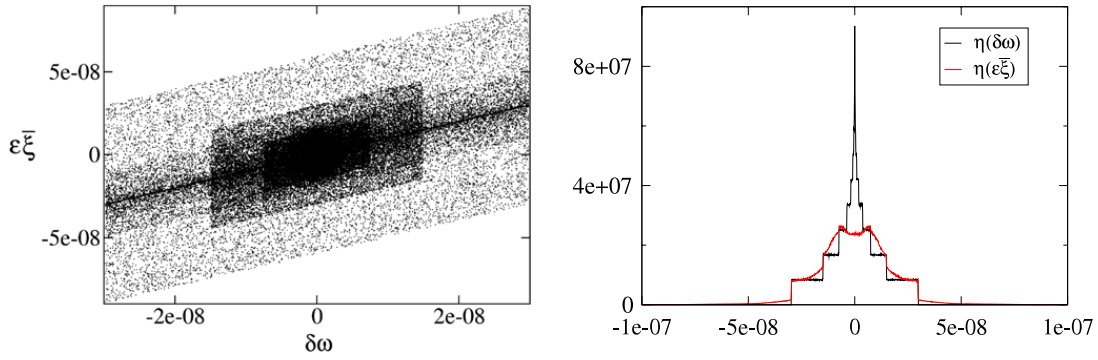

Fig. 6. Left: $\bar{\xi}(\delta\omega)$ distribution (translation on the torus). Right: comparison between the $\bar{\xi}$ (red) and $\delta\omega$ (black) distributions. Both figures are obtained with a Monte Carlo sampling over ω (10^5 values for the first figure, 10^7 for the second one). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where $\chi_I(x)$ denotes the characteristic function of the interval I, is stepwise as a consequence of the exponent in the finite-precision representations of reals, with accuracy $\epsilon = 2^{-p}$. The stepwise shape of $\eta(\delta\omega)$ is clearly also reflected in the distribution law for $\bar{\xi}$, as shown in Fig. 6. These results suggest that an uniform integration over ω implies an integration over $\bar{\xi}$. This integration over $\bar{\xi}$ should correspond more or less to a random or pseudo-random process, since is related to a truncation of the less significant digits. The probability distribution for $\bar{\xi}$ following an uniform integration for ω is not uniform, as shown in Fig. 6, but in the following discussions we will often approximate it with a uniform one.

The presence of a non-zero average contribution of noise (15) leads to a quite different situation with respect to that obtained with random white noise, in which this contribution is not present.

Even if these results do not show a qualitative dependence on the hardware architecture used to implement the algorithm (while a quantitative dependence can be found; see the right frame of Fig. 7), they can depend strongly on the algorithm. The results just shown, for which the period of w_n is of order 100 and its magnitude of order 10, were obtained by implementing the map as a translation on the torus $x' = x + \omega \mod 1$. From an analytical point of view, this is equivalent to a

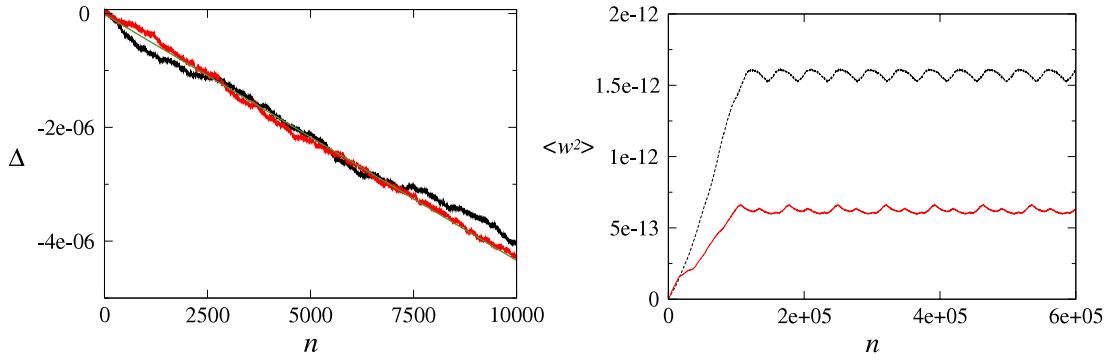

Fig. 7. Left: $\Delta_*^n(x,\omega)$ for two different, randomly chosen, x values (black and red), compared with an average over 10^4 values of x (green). $\omega=\sqrt{2}/5$. Map realized as a 2D rotation. Right: variance of w_n for $\omega=\sqrt{2}/10$, average over 10^4 different initial conditions, map realized as a 2D rotation. Black and dashed corresponds to the results obtained on the Darwin architecture, while red and continuous to those obtained on the IRIX architecture. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

two-dimensional (2D) rotation on a circle, but from a numerical point of view the latter leads to a sequence w_n which has initially, for a number of iterations of order 10^5, properties similar to a random sequence (i.e., its variance grows approximately with a linear law); see Fig. 7. For the translation on the torus, $\overline{\xi}(\omega)$ is given by $\langle\Delta_*^n(x,\omega)\rangle_x/n$ with a nice accuracy even for low values of n.

Concerning the decay of fidelity, for regular maps we distinguish between isochronous maps, for which the frequency is fixed, such as translations on the torus $x'=x+\omega \bmod 1$, and anisochronous maps, such as the skew map on the cylinder $x'=x+\omega(y)$, $y'=y$, for which the frequency depends on the initial conditions. In the former case, the fidelity does not decay, while in the latter it has a power-law decay. For isochronous maps the numerical realization of the system is basically equivalent to a deterministic rotation with a frequency $\omega+\varepsilon\overline{\xi}$, and it can be easily shown that in this case the fidelity does not decay, but oscillates with constant amplitude. Indeed, ignoring the contribution of the term w_n, we can write the action of n iterations of the numerical map as $x_n^*=x_0+n(\omega+\varepsilon\overline{\xi})$. Substituting in the Fourier analysis of Appendix B in [12] and keeping $\overline{\xi}$ fixed (i.e., not performing the integral in the noise), this leads to an oscillating phase term but not to a decay term (actually the presence of the oscillating term w_n can cause a decay of fidelity for the highest Fourier modes, those corresponding to distances of the order of the magnitude of w_n).

For the anisochronous map with $\omega(y)=y$ (to which any map with monotonic $\omega(y)$ can be reduced), a power-law decay for δF is actually observed. Indeed, the map in this case is $T(x,y)$ defined on the cylinder $\mathbb{T}\times[a,b]$ or the torus $\mathbb{T}^2$ according to

$$T_x=x+y \bmod 1 \qquad T_y=y,$$

and the fidelity with respect to the perturbed map is computed by integrating over x, y. If T_* denotes the map with numerical noise, the map becomes $T_*(x,y)$, whose iterate can be written

$$x_n^*=x_0+\epsilon\xi_x(x_0)+n(y_0+\epsilon\xi_y(y_0)) \qquad y_n^*=y_0+\epsilon\xi_y(y_0).$$

In writing this equation, we use the previous results for translations on the torus, where the role of $\overline{\xi}$ is now played by ξ_y and that of ω by y_0. We have already ignored the bounded term w_n but we have explicitly written the round-off error of x_0 for symmetry with respect to y_0. Let us consider Ψ and Φ to be functions of the only variable x, and use the usual Fourier analysis approach. The integrals can be computed if we replace, again following the previous discussion, $\xi_x(x_0)$ and $\xi_y(y_0)$ with $\overline{\xi}_x$ and $\overline{\xi}_y$, assuming that they are random variables ranging in $[-1,1]$ with uniform distribution so that the integration over x_0, y_0 is replaced by an integration over x_0, y_0 and $\overline{\xi}_x$, $\overline{\xi}_y$. The fidelity F_n is given by

$$F_*^n=\sum_{k\in\mathbb{Z}}\Phi_k\Psi_{-k}\frac{\sin(2\pi nk\varepsilon)}{2\pi nk\varepsilon}\frac{\sin(2\pi k\varepsilon)}{2\pi k\varepsilon}. \tag{16}$$

This rather strong assumption provides a result which is in good agreement with the numerical computations; see Fig. 8. A map where at any step both x and y were affected by a random error

$$T_{\epsilon x}=x+y+\epsilon\xi_x \qquad T_{\epsilon y}=y+\epsilon\xi_y$$

would give a completely different result. Indeed, taking into account that the iterate of order n is

$$x_n=x_0+ny_0+\varepsilon\left(\sum_{j=1}^{n}\xi_x^j+\sum_{j=1}^{n-1}(n-j)\xi_y^j\right) \qquad y_n=y_0+\epsilon\sum_{j=1}^{n}\xi_y^j,$$

the following result for the fidelity is found:

$$F_*^n=\sum_{k\in\mathbb{Z}}\Phi_k\Psi_{-k}\left(\frac{\sin(2\pi k\varepsilon)}{2\pi k\varepsilon}\right)^n\prod_{j=1}^{n}\frac{\sin(2\pi jk\varepsilon)}{2\pi jk\varepsilon}. \tag{17}$$

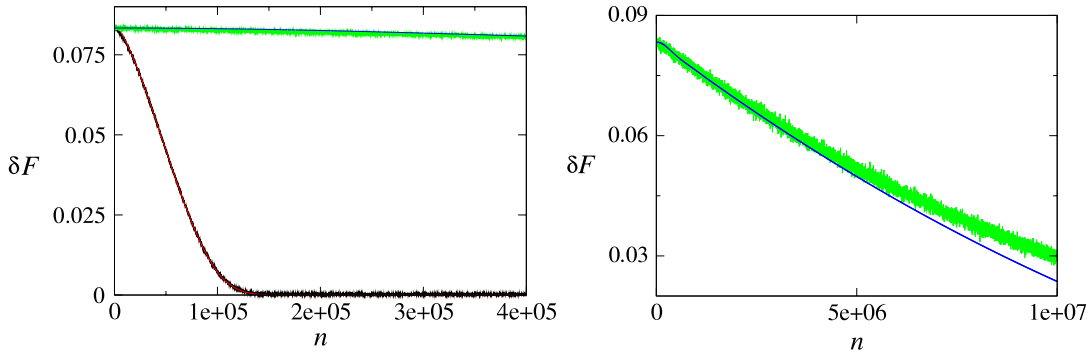

Fig. 8. Left: Decay of fidelity for the skew map. Black: additive noise, compared with its analytical prediction equation (17) (red); green: round-off noise compared with its analytical prediction equation (16) (blue). Right: comparison of round-off noise and analytical prediction on a longer time scale. $N = 10^4$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

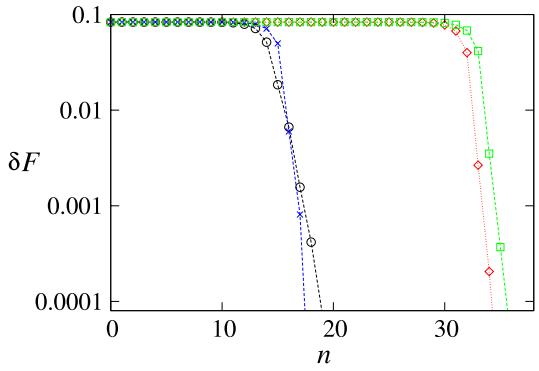

Fig. 9. Decay of fidelity for $3x \bmod 1$ represented in single precision (black, circles) and double precision (red, diamonds) compared to a reference map $T_\dagger$ using 32 digits; the results are compared to the decay of fidelity for random (additive) noise with $\varepsilon = 2^{-25}$ (blue, crosses) and $\varepsilon = 2^{-53}$ (green, squares), the value of single-precision and double-precision last significant bits for reals on the torus. Monte Carlo integrals with $N = 10^7$.

Fig. 8 shows that this result agrees with the fidelity computed numerically for a randomly perturbed map and differs drastically from the fidelity of the map with round-off noise.

6. Chaotic maps

We analyse in this section the behavior of chaotic maps, also comparing the results of numerical noise with those obtained using *random quenched noise*.

6.1. $qx \bmod 1$

The fidelity decay for the numerical map $qx \bmod 1$ shows at first sight a good qualitative agreement with the results obtained for additive noise. In particular, also for numerical maps, (i.e., for the fidelity computed comparing the "exact" map with a map perturbed by round-off) we see the presence of a threshold below which fidelity is constant and the error function is qualitatively a δ function (its support is many orders of magnitude smaller than the size of the phase space, as happened in Fig. 2). Beyond this threshold, which we can call n_*, and which grows as $-\ln \varepsilon$, i.e., linearly in the number of bits used to represent real numbers (Fig. 9), the error distribution spreads quickly over its accessible range, as can also be checked using a Monte Carlo sampling of the error distribution. These results do not show a significant dependence on the choice of the architecture.

Nevertheless, a more detailed analysis shows some important differences with additive noise, which lead us to the conclusion that our *ansatz*, according to which the sequence of local errors ξ_n at step n (see Eq. (8)) can be considered equivalent to a random one, does not apply to this map. For example, from a detailed observation of Fig. 9 is possible to see, in particular for single precision, that some difference between round-off noise and additive noise is present. In order to better understand the nature of this map under the action of round-off noise, we can study the distribution of errors ξ_n at different steps n. In the case of additive noise we would have a continuous uniform distribution identical at each step n. As a consequence, our assumption on the equivalence of round-off and additive random errors will be less and less valid as the difference of our distribution from a continuous uniform one increases. For the map $3x \bmod 1$, we have empirically found

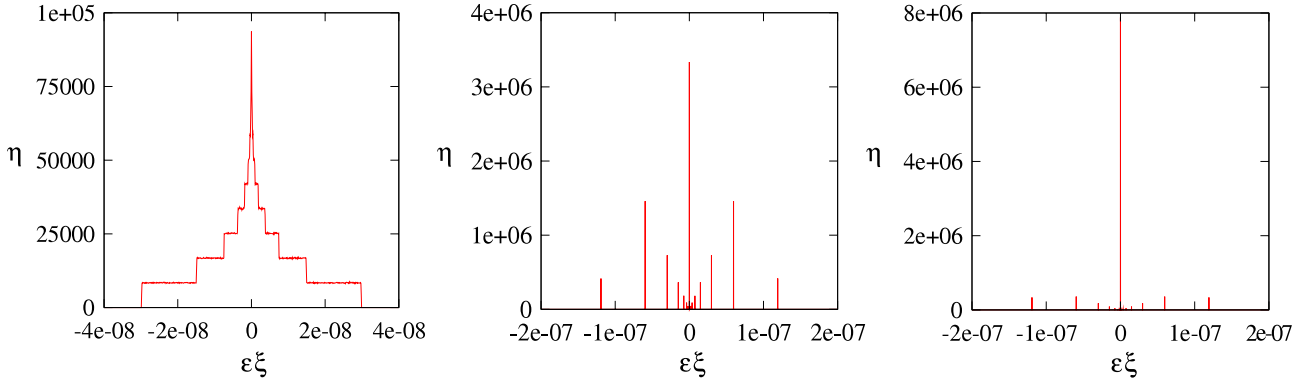

Fig. 10. Single-step error distribution for the map $3x \bmod 1$ in single precision. Left: ξ_0; center ξ_1; right ξ_2. Monte Carlo sampling with $N = 10^7$.

that the global error can be described as

$$\Delta_*^n(x) = \sum_{i=0}^{n} 3^{n-i} \varepsilon \xi_i(x) \tag{18}$$

(this equation, as Eq. (14), does not take in account boundary effects), where the round-off $\epsilon\xi_0$ on the initial condition x_0 has a step-wise continuum spectrum

$$\eta(\varepsilon\xi_0) = 2^{p-1} \sum_{k=p}^{\infty} \chi_{[-2^{-k},2^{-k}]}(\varepsilon\xi_0)$$

once sampled over the space of initial configurations, due to the exponent in the floating-point representation, where 2^{-p} is the value of the least significant bit, so that we take $\varepsilon = 2^{-p}$. The spectrum of the ξ_n for $n = 1$ is discrete and is almost completely reduced to zero for $n \geq 2$. This means that no relevant errors are made after the first iteration (Fig. 10). Due to this effect, which is caused by the extremely simple algorithmic nature of the map, the sequence of single-step errors cannot be used as a representative sequence for an integral over the noise.

It is then evident that only the initial round-off has a continuous (even if not uniform) distribution, and can be in some way considered as equivalent to the additive noise. In the discussion of the skew map, we have seen that the initial round-off determined the decay law of fidelity, and this is true also for the map $3x \bmod 1$. Let us assume for simplicity's sake, as we have previously done for the skew map, that the integral on the initial condition x^0 implies a uniform integral on the initial round-off error, and let us ignore the contribution of the rest of the single-step error sequence, ξ_n with $n \geq 1$. It is possible to show using Fourier analysis that the decay law for the fidelity of a $qx \bmod 1$ map perturbed only at the initial step, with an uniform integral over noise is

$$\delta F_\varepsilon^n = \sum_{k\neq 0} \Phi_k \Psi_{-k} \frac{\sin(2\pi\varepsilon k q^n)}{2\pi\varepsilon k q^n} \tag{19}$$

(this result is trivially obtained following the Fourier analysis procedure in Appendix B of [12] and performing only the first integral over noise). The decay law (19) describes the decay law for fidelity due to round-off noise better than (6), showing that the main contribution to decay fidelity is due to the initial, continuous spectrum perturbation (Fig. 11). To analyse the contribution of ξ_n with $n \geq 1$, we remove the initial round-off. To this end, we choose rational initial conditions that can be represented exactly with the prescribed precision. In this case, the fidelity presents a threshold, but after the threshold oscillates without decaying to zero (Fig. 12).

After the threshold, the error distribution expands to the whole phase space but does not converge to an asymptotic distribution; see Fig. 13.

This effects are due, according to us, to the extremely simple algorithmic nature of this map.

6.2. Other chaotic maps

For other, (slightly) more algorithmically complex maps, such as the Bernoulli, Hénon, Logistic, Intermittent and Baker maps and the Standard map with $K \gg 1$, the sequence of errors ξ_i has a continuous spectrum for each value of i, and converges quickly to an asymptotic distribution (Fig. 14). For all these systems the asymptotic error distribution results in being almost indistinguishable from that obtained using additive noise (Fig. 15), and the decay law is almost identical to the one obtained for additive noise (Fig. 16), suggesting that for these maps the assumption of equivalence between round-off noise and additive noise could be valid (for these maps the initial round-off error plays no special role). These results do not depend on the architecture.

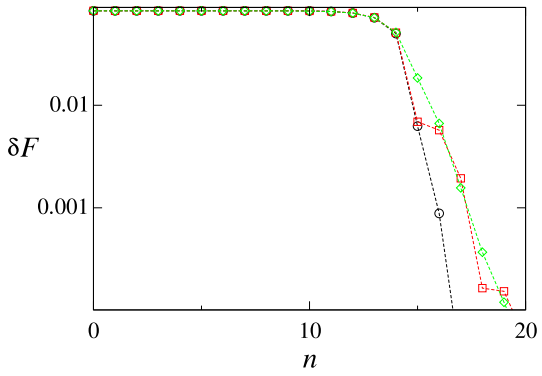

Fig. 11. Decay of fidelity for $3x \bmod 1$. Black, circles: analytical result (6), using $\varepsilon = 2^{-25}$. Red, squares: analytical result (19). Green, diamonds: round-off noise (single precision compared to double precision), $N = 10^7$.

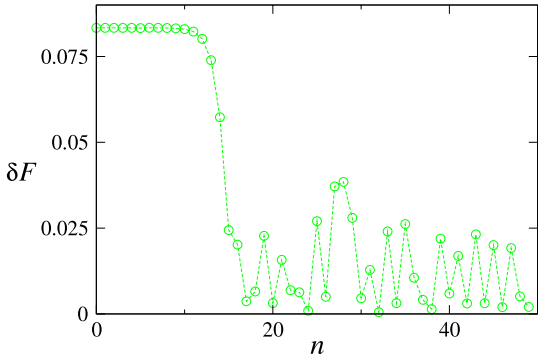

Fig. 12. Decay of fidelity for $3x \bmod 1$, obtained comparing a single-precision map with a double-precision one (round-off noise) using initial conditions that can be exactly represented in single precision, i.e., removing the initial round-off. Monte Carlo integrals, $N = 10^6$.

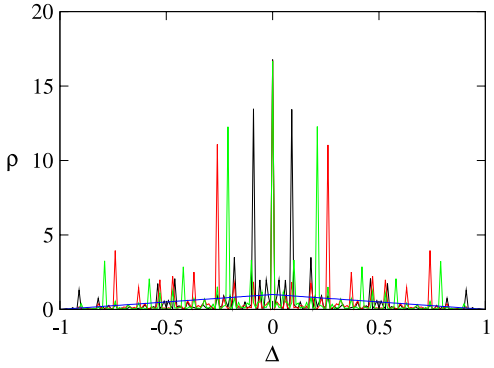

Fig. 13. Error distribution ρ^n obtained comparing a single-precision $3x \bmod 1$ map with a double-precision map, using initial conditions in single precision (no initial round-off). Black, $n = 30$; red, $n = 31$; green $n = 32$; compared to the triangular function (blue) obtained using initial condition in double precision (initial round off). Monte Carlo samplings using $N = 10^7$ initial conditions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

For all chaotic maps it is then possible to identify a threshold n_*, which grows as $-\ln \varepsilon$, i.e., linearly in the number of bits used to represent real numbers, beyond which the error distribution spreads quickly to the whole phase space, and which can be interpreted as a limit for the reliability of numerical computations (at least for what concerns punctual results).

6.3. Quenched noise

It is interesting to study the decay of fidelity for maps perturbed with *quenched noise*, i.e., without performing the integral over noise. In this approach a single succession ξ_i of single-step errors is randomly chosen and applied to all initial conditions. This case is interesting because it shows the difference between performing or not performing the integral over noise and thus can clarify the relation between numerical and additive noise.

For the Bernoulli, Hénon, Logistic, Intermittent and Baker maps and the Standard map with $K \gg 1$, quenched noise results in being almost equivalent to additive and numerical noise (just a little more irregular due to the absence of the integral over noise; see Fig. 17). For the map $3x \bmod 1$, it can be shown (generalizing a result in [17]) that the fidelity

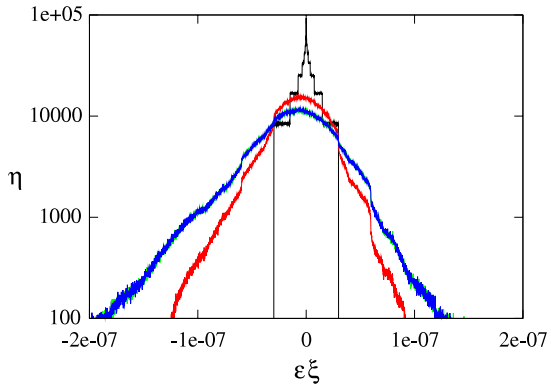

Fig. 14. Distribution of single-step errors ξ_i for the Hénon map, x observable, $N = 10^7$. Black, $i = 0$. Red, $i = 1$. Green, $i = 20$. Blue $i = 50$ (green and blue are very difficult to distinguish). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

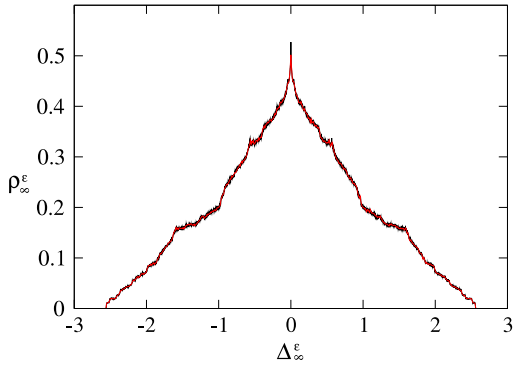

Fig. 15. Asymptotic error distribution for the Hénon map, x observable. Black, additive noise ($\varepsilon = 10^{-8}$), red comparison between single and double precision (the two functions are almost identical). Monte Carlo samplings using $N = 10^7$ initial conditions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

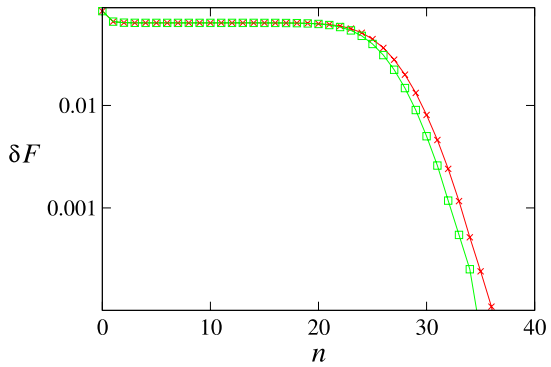

Fig. 16. Decay of fidelity for the Baker map. Green, squares: single-precision round-off noise, showing the same decay law of additive noise with $\varepsilon = 2^{-25}$ (red, crosses). Monte Carlo integrals with $N = 10^7$.

for quenched noise oscillates but does not decay to zero. We have found that this result is particularly difficult to verify using Monte Carlo numerical integration. As is shown in Fig. 18, the map $3x \mod 1$, studied with numerical precision ε_2 and perturbed with quenched noise $\varepsilon_1 \gg \varepsilon_2$, starts oscillating in correspondence with the fidelity threshold for ε_1 (in agreement with the analytical result) but then drops to zero in correspondence with the fidelity threshold for ε_2 (see Fig. 18, comparing with Fig. 9 to know the single-precision and double-precision thresholds). According to us, this result supports our conclusion that the fidelity decay threshold is a good measure for the maximum number of iterations under which the numerical computations are completely reliable. Indeed, as long as the round-off noise is negligible, the fidelity in the presence of quenched noise oscillates rather than decays, as theoretically established, but later on when the round-off noise becomes effective and the numerical computations are no longer reliable a decay is observed. Not taking the round-off into account would provide a result that contradicts rigorous theoretical results.

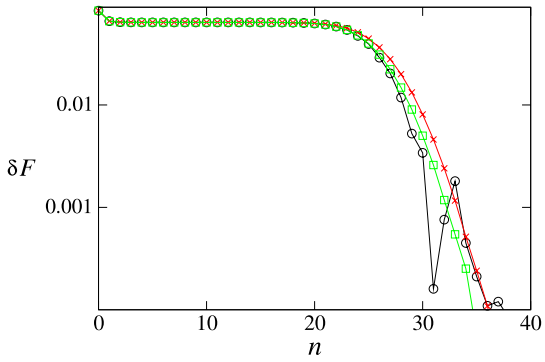

Fig. 17. Decay of fidelity for the Baker map. Black, circles: the sequence errors of order 2^{-25} is the same for each initial condition (quenched noise), showing a more irregular decay law; green, squares: single-precision round-off noise, showing the same smooth decay of additive noise with $\varepsilon = 2^{-25}$ (red, crosses). Monte Carlo integrals with $N = 10^7$.

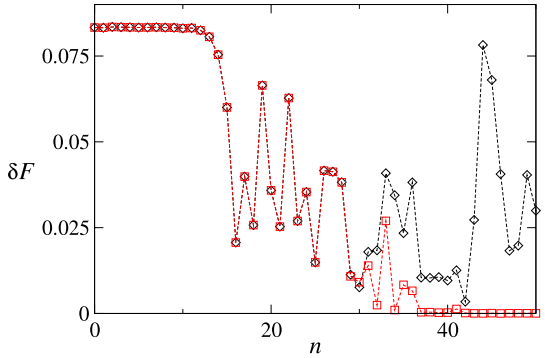

Fig. 18. Decay of fidelity for $3x$ mod 1. Black, diamonds: quenched noise computed with 32 digits. Red, squares: quenched noise computed with double precision. Magnitude of quenched noise $\varepsilon = 2^{-25}$ for both figures. $N = 10^6$.

7. Conclusions

We have used the results of a previous work on additive noise to study the effects of round-off noise on discrete systems. For regular maps, the behavior depends on the algorithmic realization and on its character: for isochronous maps, the fidelity error does not decay, for anisochronous maps it has a power-law decay, whereas in the presence of an additive noise the decay is always exponential. For chaotic systems, the round-off noise and additive noise are almost equivalent. We also showed that the fidelity is an efficient tool for the study of the time scales for the convergence to the asymptotic error distribution, since, at least for chaotic maps, it allows us to find a threshold value below which the numerical system can be considered as equivalent to the exact one. This threshold linearly grows with the number of bits used to represent real numbers.

Acknowledgements

G. Turchetti acknowledges financial support from the PRIN 2007 grant *Simulazione della produzione di elettroni dalla interazione laser-(pre)plasma nel contesto della ignizione veloce* coordinated by D. Batani. Francesco Zanlungo has been partially supported by the GDRE 224 GREFI-MEFI (jointly sponsored by the French CNRS and the Italian INDAM).

References

[1] A. Katok, B. Hasselblatt, Introduction to the Modern Theory of Dynamical Systems, Cambridge University Press, Cambridge, 1995.
[2] C. Grebogi, S. Hammel, J. Yorke, J. Complexity 3 (1987) 136.
[3] S.M. Hammel, J.A. Yorke, C. Grebogi, Bull. Amer. Math. Soc. 19 (1988) 465.
[4] C. Grebogi, S. Hammel, J. Yorke, T. Sauer, Phys. Rev. Lett. 65 (1990) 1527.
[5] S.N. Chow, K. Palmer, J. Dynam. Differential Equations 3 (1991) 361.
[6] E.S. Van Vleck, SIAM J. Sci. Comput. (USA) 16 (1995) 1177.
[7] T. Sauer, C. Grebogi, J. Yorke, Phys. Rev. Lett. 79 (1997) 59.
[8] T. Sauer, Phys. Rev. E 65 (2002) 036220.
[9] A. Crisanti, M. Falcioni, G. Mantica, A. Vulpiani, Phys. Rev. E 50 (1994) 1959–1967.
[10] M. Falcioni, A. Vulpiani, G. Mantica, S. Pigolotti, Phys. Rev. Lett 91 (2003) 44101.
[11] G. Mantica, Phys. Rev. E 61 (2000) 6434–6443.
[12] P. Marie, G. Turchetti, S. Vaienti, F. Zanlungo, Chaos 19 (2009) 043118.
[13] G. Turchetti, S. Vaienti, F. Zanlungo, Europhys. Lett. 89 (2010) 40006.
[14] Y. Kifer, Ergodic Theory of Random Transformations, Birkhäuser, 1986.
[15] D. Knuth, The Art of Computer Programming, vol. 2, Addison-Wesley, 1969.
[16] G. Benenti, G. Casati, G. Veble, Phys. Rev. E 670 (2003) 055202-1.
[17] C. Liverani, Ph. Marie, S. Vaienti, Random classical fidelity, J. Stat. Phys. 128 (4) (2007) 1079.