Article

The Emergence of the Normal Distribution in Deterministic Chaotic Maps

Damián H. Zanette [1,2,*,†] and Inés Samengo [1,2,*,†]

1 Instituto Balseiro, Centro Atómico Bariloche Comisión Nacional de Energía Atómica, Universidad Nacional de Cuyo, Av. Ezequiel Bustillo 9500, San Carlos de Bariloche 8400, Argentina
2 Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina
* Correspondence: damian.zanette@ib.edu.ar (D.H.Z.); ines.samengo@ib.edu.ar (I.S.)
† These authors contributed equally to this work.

Abstract: The central limit theorem states that, in the limits of a large number of terms, an appropriately scaled sum of independent random variables yields another random variable whose probability distribution tends to attain a stable distribution. The condition of independence, however, only holds in real systems as an approximation. To extend the theorem to more general situations, previous studies have derived a version of the central limit theorem that also holds for variables that are not independent. Here, we present numerical results that characterize how convergence is attained when the variables being summed are deterministically related to one another through the recurrent application of an ergodic mapping. In all the explored cases, the convergence to the limit distribution is slower than for random sampling. Yet, the speed at which convergence is attained varies substantially from system to system, and these variations imply differences in the way information about the deterministic nature of the dynamics is progressively lost as the number of summands increases. Some of the identified factors in shaping the convergence process are the strength of mixing induced by the mapping and the shape of the marginal distribution of each variable, most particularly, the presence of divergences or fat tails.

Keywords: stable distributions; deterministic systems; central limit theorem

1. Introduction

According to the central limit theorem (CLT), the sum of independent variables with finite first and second moments is governed by a Gaussian distribution when the number of summands is asymptotically large. The mean value and the variance of the Gaussian equal the sum of the individual mean values and variances, respectively. The Gaussian distribution has maximal entropy for a given variance and is reached independently of the distributions from which the summands are sampled. The convergence to the Gaussian limit, therefore, can be viewed as a loss of information about the original data. Extension to sums of variables with diverging first and second moments have been derived in [1–4]; the asymptotic distributions of there are no longer Gaussian, but are still members of a family of so-called *stable distributions*.

Experience shows that many systems are successfully modeled by stable distributions, for example, in the theory of errors and propagation of uncertainty. This is often justified by the fact that errors, as well as many other quantities of interest, can be conceived as the sum of a large number of variables representing disparate magnitudes that appear to be unrelated. Yet, Physics, for instance, dictates that all the variables describing a system of interacting particles (as opposed to an ensemble of free particles) are correlated to one another. Therefore, the independence condition is no more than an approximation.

To improve this approximation, extensions of the CLT have also been developed for variables that bear different degrees of statistical dependencies, including those obtained through the subsequent application of a deterministic rule that produces ergodicity and

aperiodicity [5–11]. Here, we analyze several systems of this type. As discussed in the next section, a conveniently modified version of the CLT exists for appropriately scaled sums of variables deterministically related to one another. Notably, the family of stable distributions for these cases coincides with the ones obtained for independent variables. These extensions provide mathematical certainty that sums of strongly correlated variables, if produced through a chaotic dynamical system, lose all memory of their original distribution, and asymptotically approach a distribution that also happens to be the limit of sums of independent variables sampled from a certain distribution. The strong statistical dependencies governing the physical world, therefore, may be legitimately ignored when describing the probability distributions of macroscopic variables, and it is legitimate to conceive the latter as a sum of a large number of microscopic, independent variables. This property greatly simplifies the description of macroscopic systems and has probably played a crucial role in the development of the theory of probability.

In practical situations, however, it is important to know how many terms a sum needs to include for its distribution to be well described by the asymptotic result. To shed light on this question, we study in this paper the convergence of the probability distribution of a sum of perfectly correlated variables, generated through the iteration of a chaotic, deterministic map, towards the asymptotic distribution predicted using the extensions of the CLT. The aim is to characterize how the loss of information about the deterministic nature of the map depends on the number of variables that are summed together. Since previous theoretical results do not predict the rate of convergence towards asymptotic distributions in deterministic systems, our analysis is based on numerical simulations of several paradigmatic examples, and on a comparison with the behavior of randomly sampled systems with the same distributions.

The paper is organized as follows. In Section 2, we present the main theoretical tools to be employed later; these include the extension of the CLT to variables that are strongly correlated, the information–theoretical measures that quantify the differences between probability distributions, and the behavior of the variance of a sum of variables that are correlated. The following three sections apply these tools to the analysis of a chaotic dynamical system with a uniform marginal distribution and varying Lyapunov exponent (the Bernoulli map, Section 3), a chaotic dynamical system with a highly nonuniform marginal distribution and several types of orbits (the logistic map, Section 4), and an example of a process with fat-tailed distribution (Section 5). Our main conclusions are summarized in Section 6.

2. Central Limit Theorem for Deterministic Maps

We consider a generic one-dimensional map, $x(t+1) = f[x(t)]$, with a well-defined invariant measure $\rho_x(x)$, determined using the identity

$$\rho_x(x)dx = \rho_x[f(x)]df(x) = [\rho_x \circ f](x)\, f'(x)dx, \tag{1}$$

where $[\rho_x \circ f](x)$ is the composition of functions $\rho_x(x)$ and $f(x)$, and the prime indicates differentiation with respect to x. We assume that the mean value $\bar{x}$ is finite over the distribution $\rho_x(x)$, and—for now—we assume the variance σ_x^2 of x is also finite:

$$\bar{x} = \int x\rho_x(x)dx < \infty, \quad \sigma_x^2 = \int (x-\bar{x})^2\rho_x(x)dx < \infty, \tag{2}$$

where the integrals run over the whole domain of variation of x. In Section 5, we study a case where we relax the condition that σ_x^2 is finite. A central limit theorem (CLT) for this kind of system applies [6–11] when the map under study is ergodic and aperiodic. We recall that a map is ergodic if all its invariant sets are null or co-null, and it is aperiodic if its periodic orbits form a null set [11]. The combination of ergodicity and aperiodicity is typically equivalent to the dynamics being chaotic [12]. In this case, the CLT states that the distribution of the (centered, suitably normalized) sums of N successive values of $x(t)$,

$$s_N(t) = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} [x(t+k-1) - \bar{x}], \tag{3}$$

becomes normal for $N \to \infty$:

$$\rho_s(s) = \frac{1}{\sqrt{2\pi\sigma_s^2}} \exp\left(-\frac{s^2}{2\sigma_s^2}\right) \equiv G_{\sigma_s}(s), \tag{4}$$

for some value of the variance σ_s^2. Here, G_{σ_s} denotes the Gaussian centered at zero, with standard deviation σ_s.

For each value of N, the variables $x(t)$ and $s_N(t)$ can be integrated into a single two-dimensional map:

$$\begin{cases} x(t+1) = f[x(t)], \\ s_N(t+1) = s_N(t) - \frac{1}{\sqrt{N}} x(t) + \frac{1}{\sqrt{N}} f^{(N)}[x(t)], \end{cases} \tag{5}$$

where $f^{(N)}(x) = \overbrace{f \circ f \circ \cdots \circ f}^{N}(x)$ is the N-th self-composition of $f(x)$. Thus, for $N \to \infty$, the marginal invariant measures of the variables x and s_N in map (5) are, respectively, $\rho_x(x)$ and the Gaussian $\rho_s(s) = G_{\sigma_s}(s)$ of Equation (4).

In contrast with the sums of statistically independent random variables drawn from a given distribution, in the limit $N \to \infty$, the variance of the sums $s_N(t)$ does not necessarily coincide with that of the summands, σ_x^2. The difference arises from the correlations between successive values of $x(t)$, induced by the map $x(t+1) = f[x(t)]$, with the ensuing mutual correlations between the values of $s_N(t)$. For a finite number of summands N, the variance of $s_N(t)$ is given by the Green–Kubo formula [13]:

$$\sigma_{s_N}^2 = \sigma_x^2 + 2 \sum_{k=1}^{N-1} \left(1 - \frac{k}{N}\right) \overline{[x(t) - \bar{x}][x(t+k) - \bar{x}]}, \tag{6}$$

where the overline indicates the average with respect to the distribution $\rho_x(x)$. The value of $\sigma_{s_N}^2$ becomes independent of t when the process $x(t)$ has reached a stationary regime. For $N \to \infty$, the variance is

$$\sigma_s^2 = \lim_{N \to \infty} \sigma_{s_N}^2 = \sigma_x^2 + 2 \sum_{k=1}^{\infty} \overline{[x(t) - \bar{x}][x(t+k) - \bar{x}]}. \tag{7}$$

Provided that the sum converges, this formula gives the variance of the asymptotic normal distribution $G_{\sigma_s}(s)$ of increasingly long sums $s_N(t)$.

In the following, we study the process of convergence towards the asymptotic distribution predicted by the above CLT for some selected deterministic maps, as the number of terms in the sums s_N grows. For each N, we numerically iterate Equation (5) and estimate the distribution of the sums s_N, $\rho_{s_N}(s_N)$, as a suitably normalized 10^3-column histogram built from, typically, 10^7 values of s_N. To quantify the difference between ρ_{s_N} and the expected asymptotic Gaussian distribution $G_{\sigma_s}(s)$, we use the Kullback–Leibler divergence (KLD). Recall that the KLD between two distributions $\rho_1(s)$ and $\rho_2(s)$ is defined as

$$D(\rho_1 || \rho_2) = \int \rho_1(s) \log_2\left[\frac{\rho_1(s)}{\rho_2(s)}\right] ds. \tag{8}$$

This quantity measures the inefficiency with which the data s is represented by a code optimized to be maximally compact under the assumption that the distribution is ρ_2 when, in reality, the data are generated from ρ_1. The inefficiency equals the mean number of extra bits per sample [14]. The divergence only vanishes when the two distributions coincide, and is otherwise positive. For brevity, we hereafter denote as D_G the KLD between the distribution ρ_{s_N} and the asymptotic normal distribution G_{σ_s}: $D_G \equiv D(\rho_{s_N} || G_{\sigma_s})$.

Additionally, for each N, it is interesting to compare ρ_{s_N} with a normal distribution with the variance $\sigma_{s_N}^2$ given by Equation (6), namely, the same variance as the sums s_N. Since $\sigma_{s_N}^2 \to \sigma_s^2$ as $N \to \infty$, this is an alternative way of characterizing the convergence to the asymptotic Gaussian G_{σ_s}. For this comparison, we introduce the KLD $D_{G_N} \equiv D\left(\rho_{s_N} || G_{\sigma_{s_N}}\right)$.

Finally, in order to contrast the deterministic dynamics of the chaotic map under study with a genuinely aleatory process, we calculate the KLD for the distribution of sums of the same form as in Equation (3), but with the N values of the variable x drawn at random from the invariant measure $\rho_x(x)$. According to the standard CLT for statistically independent variables, as N grows, the distribution $\rho_{s_N}^{\text{random}}$ of these random sampling sums is expected to asymptotically converge to a Gaussian with variance σ_x^2. To quantify this convergence, we compute $D_{\text{random}} \equiv D\left(\rho_{s_N}^{\text{random}} || G_{\sigma_x}\right)$.

The measures D_{random}, D_G and D_{G_N} reflect three different aspects of the convergence of ρ_{s_N} to G_{σ_s}. The process by which D_{random} tends to zero describes how independent variables, when summed together, lose the memory of the distribution from which they are sampled and approach a Gaussian. The Gaussian distribution is the one with maximal entropy among those with fixed variance. By acquiring a Gaussian shape, therefore, the distribution of the sum maximizes uncertainty. In Appendix B, we show that, for large N, the divergence D_{random} decays as N^{-1} if ρ_x is not symmetric around its mean value, and at least as fast as N^{-2} if there is symmetry.

A steep decay of D_{G_N} with N, at a faster rate than D_G, implies a rapid evolution of ρ_{s_N} towards a bell-shaped distribution, whose variance may still have to evolve to its asymptotic value σ_s^2. The convergence process can therefore be conceived as a sequence of two stages, the first one consisting of shedding all the structure in $\rho(x)$ and becoming Gaussian-like, and the second, adjusting the variance. Once ρ_{s_N} is approximately Gaussian, its KLD with the asymptotic distribution G_{σ_s} can be analytically calculated in terms of their respective variances:

$$D_G \approx \log_2\left(\frac{\sigma_{s_N}}{\sigma_s}\right) + \frac{1}{2 \ln 2} \frac{\sigma_s^2 - \sigma_{s_N}^2}{\sigma_{s_N}^2}. \tag{9}$$

3. The Bernoulli Map

As a first case of study, we take the generalized Bernoulli map

$$x(t+1) = f[x(t)] = \{m x(t)\}, \tag{10}$$

where $\{\cdot\}$ indicates a fractional part, and $m > 1$ is an integer factor. This map has been extensively studied since long ago as a paradigm of deterministic chaotic systems, due to its combination of complex behavior and analytical traceability. Its Lyapunov exponent equals $\ln m$. The invariant measure of $x(t)$ is particularly simple:

$$\rho_x(x) = \begin{cases} 1 & \text{for } x \in [0,1), \\ 0 & \text{otherwise,} \end{cases} \tag{11}$$

for all m, with $\bar{x} = 1/2$ and $\sigma_x^2 = 1/12$. We show in Appendix A that the variances of the sums s_N can be explicitly calculated:

$$\sigma_{s_N}^2 = \frac{1}{12} + \frac{1}{6(m-1)}\left(1 - \frac{m}{m-1}\frac{1 - m^{-N}}{N}\right). \tag{12}$$

Note that for $N \gg (\ln m)^{-1}$, this variance takes the approximate form

$$\sigma_{s_N}^2 \approx \frac{1}{12}\frac{m+1}{m-1}\left(1 - \frac{N_0}{N}\right), \tag{13}$$

with $N_0 = 2m/(m^2 - 1)$. For $N \to \infty$, in turn,

$$\sigma^2_{s_N} \to \sigma^2_s = \frac{1}{12}\frac{m+1}{m-1}. \tag{14}$$

We first consider the Bernoulli map for $m = 2$. Dark full lines on the left column of Figure 1 show numerical results for the distributions of the sums s_N, ρ_{s_N}, for three small values of N. Light-colored curves stand for the asymptotic Gaussian $\rho_s = G_{\sigma_s}$, and dashed curves are the Gaussians $G_{\sigma_{s_N}}$ for each N. Their respective variances, σ^2_s and $\sigma^2_{s_N}$, are given by Equation (12). On the right column, dark- and light-colored curves, respectively, show the distributions of the sums of randomly sampled values of x, $\rho^{\text{random}}_{s_N}$, calculated analytically as N-th order self-convolutions of $\rho_x(x)$, and the expected asymptotic Gaussian G_{σ_x}. A comparison of the two columns illustrates the difference between the distributions of the sums generated by map iteration on one side and by random sampling on the other. It also shows that convergence to the asymptotic distribution is faster in the latter case.
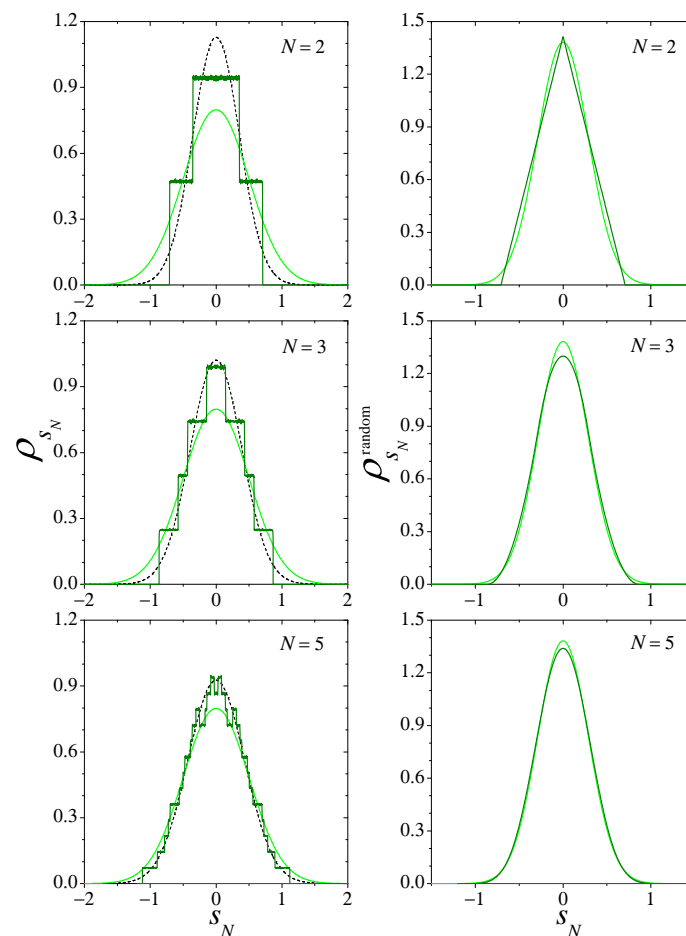

Figure 1. **Left**, dark line: Numerical results for the distribution ρ_{s_N} of the sums s_N defined in Equation (3), in the case the Bernoulli (10) map with $m = 2$, for three small values of N. The light curve is the Gaussian expected for $N \to \infty$, and the dashed curve is a Gaussian with the same variance as predicted for ρ_{s_N}. **Right**, dark curve: The distribution $\rho^{\text{random}}_{s_N}$ for sums of N values of x randomly sampled from $\rho_x(x)$ is a normalized version of the Irwin–Hall distribution [15,16], which can be obtained analytically through the successive self-convolution of ρ_x. The light curve is the Gaussian expected for $N \to \infty$. Note the different scales on the left and right columns.

The main panel of Figure 2 shows, with different symbols, the KLDs D_G, D_{G_N}, and D_{random}, defined in the preceding section. For $N = 1$, by definition, $D_{\text{random}} = D_{G_N}$. For large N (Appendix B), D_{random} decays as N^{-2}. The straight lines in the log–log plot of

the figure have slope -2, suggesting that the decay of the divergence D_G approximately follows the same asymptotic dependence on N. The inset in Figure 2 shows, as dots, the numerical estimation of the variance of s_N over the distribution ρ_{s_N} as a function of N. The dashed curve corresponds to the analytical expression of Equation (12).
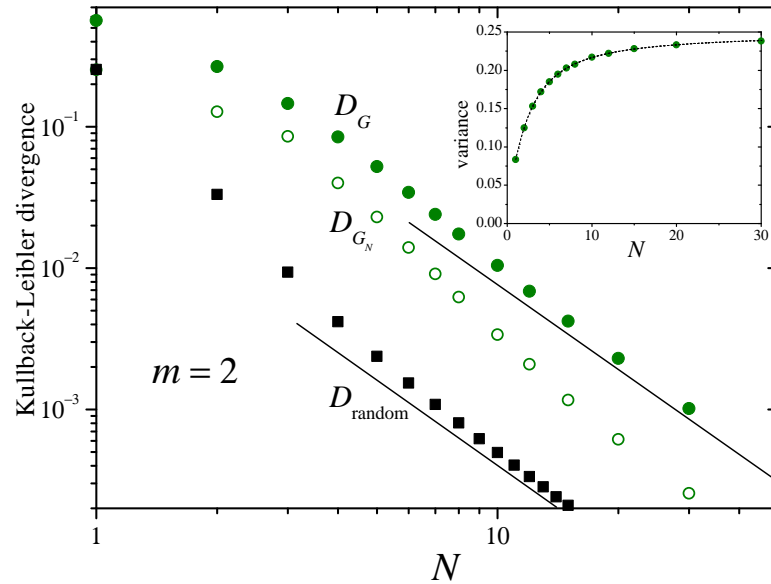

Figure 2. Main panel: The Kullback–Leibler divergences D_G, D_{G_N}, and D_{random}, defined in the text, as functions of the number of terms in the sums s_N of Equation (3), for the Bernoulli map (10) with $m = 2$. The straight lines in this log–log plot have a slope -2. The inset shows, as dots, numerical results for the variance $\sigma^2_{s_N}$ over the distribution $\rho_{s_N}(s_N)$. The dashed line joins the analytical values predicted from Equation (12).

In the range shown in the figure, for $N \gtrsim 10$, D_G is larger than D_{random} by a factor of around 14. Meanwhile, in the same range, D_{G_N} decays faster, approximately as $N^{-2.3}$. As discussed at the end of Section 2, this faster decay of D_{G_N} suggests that ρ_{s_N} is rapidly approaching a Gaussian distribution, with a KLD with the asymptotic distribution ρ_S as given by Equation (9). Replacing Equation (13) into Equation (9) and expanding up to second order in N_0/N yields

$$D_G \approx \frac{1}{4 \ln 2} \frac{N_0^2}{N^2}. \tag{15}$$

For $m = 2$ we have $N_0 = 4/3$ so that, according to the above equation, $D_G \approx 0.64 \, N^{-2}$. A power–law fitting of the data for D_G for $N \le 20 \le 50$ gives $D_G \approx 0.69 \, N^{-1.9}$, which fits the prediction of Equation (15) remarkably well. This agreement provides strong evidence in favor of the hypothesis that ρ_{s_N} converges to ρ_s in two stages, acquiring a Gaussian shape in the first, and adjusting the variance in the second. The transition from the first stage to the second, however, does not imply that ρ_{s_N} is strictly speaking a Gaussian distribution.

What are the implications of the fact that after the initial transient D_G and D_{random} both decay with the same power law, approximately proportional to N^{-2}? In this regime, $D_G \approx 14 \, D_{\text{random}}$ which means that, for each N, $D_{\text{random}}(N)$ is approximately equal to $D_G(\sqrt{14}N)$. By increasing the number of random samples drawn from the invariant measure (11), D_{random} diminishes by the same amount as D_G diminishes when running the Bernoulli deterministic mapping a re-scaled, larger number of samples, with a scaling factor of $\alpha \approx \sqrt{14} \approx 3.7$. In other words, α samples of the deterministic map are as informative about the asymptotic distribution as a single sample in the random drawing. The presence of correlations makes each new sample from the deterministic dynamics less informative (by a factor of α) than from purely independent draws.

The factor α may also be semi-quantitatively associated with the relation between the asymptotic variance σ_s^2 and the original variance σ_x^2. In Equation (3), the normalization

factor $1/\sqrt{N}$ compensates for the fact that the variance of a sum of N independent samples is proportional to N. Yet, when the summands bear statistical interdependence, the intended compensation need not be attained. The higher the correlations in the deterministic map, the less informative each new datum is, the more unsuccessful the compensation, and the larger the increase in the asymptotic variance. In the present case, the variance increases threefold, from $1/12$ to $1/4$, which is similar to the factor relating D_G and D_{random}, namely, α.

Considering now the other values of m in the Bernoulli map (10), the numerical results presented in Figure 3 show that the dependence of D_G on N is similar to that obtained for $m = 2$, with the only difference that D_G becomes progressively smaller as m grows. As before, the convergence may be conceived as consisting of two stages, with Equation (9) approximately holding for the second stage. According to the results of Figure 3, the second state is reached faster for larger values of m. As expected from the large-N asymptotic behavior of D_G predicted by Equation (15) with $N_0 = 2m/(m^2 - 1)$ [cf. Equation (13)], it approaches D_{random} for large m. This implies that the effect of the statistical dependencies induced by the deterministic nature of the map decreases as m grows. The KLD D_{G_N} is not shown in Figure 3, but its behavior is similar to that of the case of $m = 2$ (Figure 2).
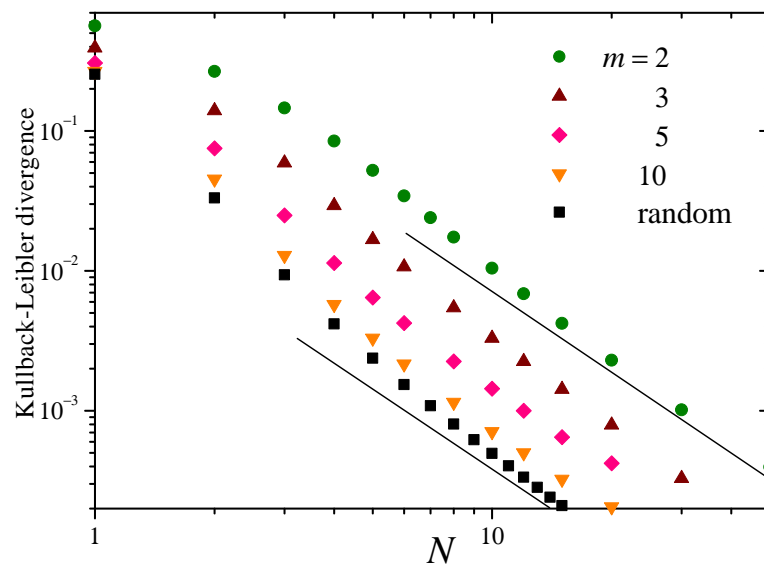

Figure 3. The Kullback–Leibler divergence D_G for the Bernoulli map (10) with various values of m, and D_{random} (which is the same for all m). The straight lines have slope -2.

In summary, in the Bernoulli map, D_{G_N} decreases faster than D_G during the first stage of the convergence process, where ρ_{s_N} acquires a Gaussian-like shape. Only later is the variance adjusted towards its final value. The second stage can be modeled analytically, providing a good qualitative description of the asymptotic behavior inferred from numerical results.

4. The Logistic Map: Full Chaos and Intermittency

We now turn our attention to the logistic map [17,18]

$$x(t + 1) = f[x(t)] = \lambda x(t)[1 - x(t)], \tag{16}$$

with $0 < \lambda \leq 4$. Much like Bernoulli's, the logistic map hardly needs any presentation. We first consider the case $\lambda = 4$, which we call the regime of "full chaos". For this value of λ, the dynamics are chaotic and therefore comply with the hypotheses of the CLT for deterministic systems discussed in Section 2. Moreover, due to the existence of a nonlinear change of variables that transforms the logistic map with $\lambda = 4$ into the Bernoulli map of Equation (10) with $m = 2$, several analytical results for the latter can be extended to

the former. In spite of this connection, as we show below, the statistics of the sums s_N are qualitatively different between the two maps.

For $\lambda = 4$, the invariant measure of the logistic map can be written explicitly as [19]

$$\rho_x(x) = \frac{1}{\pi\sqrt{x(1-x)}} \tag{17}$$

for $0 \leq x \leq 1$, and 0 otherwise. The mean value is $\bar{x} = 1/2$ and the variance is $\sigma_x^2 = 1/8$. As we show in Appendix A, the correlations between iterations of the map, $c_k = \overline{[x(t) - \bar{x}][x(t+k) - \bar{x}]}$, vanish for all k. From Equations (6) and (7), this implies that the variances of the sums s_N are the same for all N, and therefore coincide with both the variance of x and with the limit for $N \to \infty$: $\sigma_{s_N}^2 = \sigma_s^2 = \sigma_x^2$. Therefore—in contrast with the Bernoulli map studied in the preceding section—it is not possible to discern between a first stage of convergence to a Gaussian profile and a second stage of adjustment of the variance.

In Figure 4, the left column shows numerical estimations of the distributions $\rho_{s_N}(s_N)$ of the sums of N consecutive iteration of the logistic map with $\lambda = 4$, for three values of N. The light-colored curve corresponds to the expected asymptotic Gaussian.
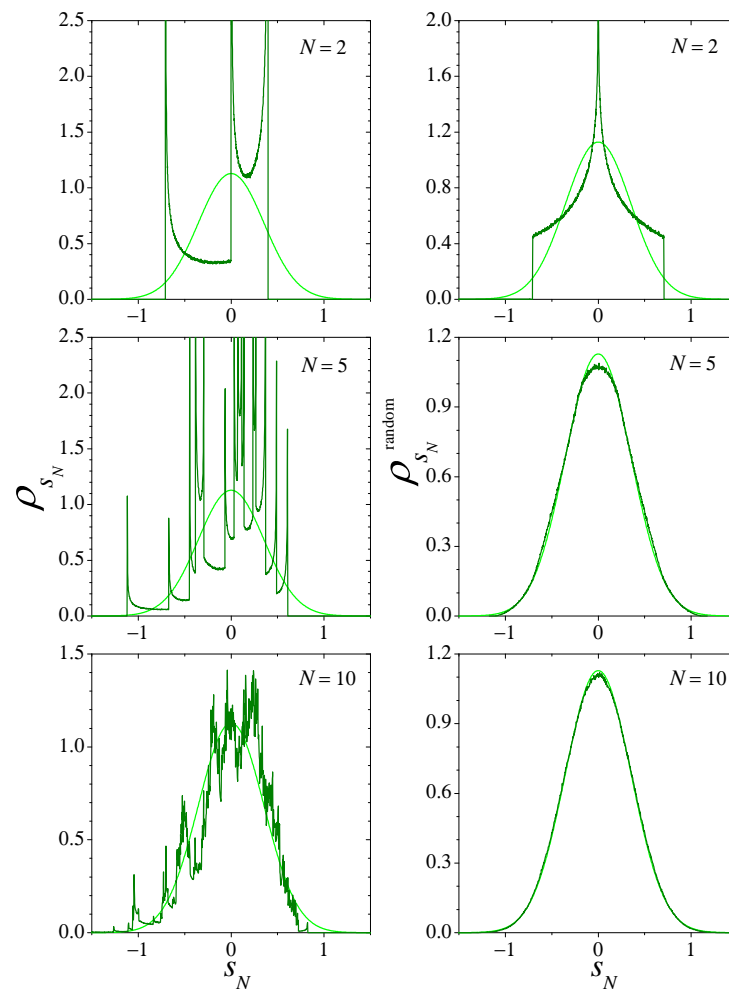

Figure 4. As in Figure 1 for the logistic map of Equation (16) in the regime of full chaos, $\lambda = 4$. The distributions $\rho_{s_N}^{\text{random}}$, dark lines on the right column, have now been obtained numerically. Note the different scales in different panels.

In addition to the sharp peaks in the profile of ρ_{s_N} for small N, an important difference with the Bernoulli map (Figure 1) is that ρ_{s_N} is no longer symmetric with respect to zero. This asymmetry may come as a surprise, taking into account that both $f(x)$ and $\rho_x(x)$ are

symmetric around the mean value $\bar{x}$. The asymmetry, however, originates from the fact that the functions $x + f(x)$, $x + f(x) + f^{(2)}(x)$, $x + f(x) + f^{(2)}(x) + f^{(3)}(x)$, ..., which ultimately determine the distributions of the sums s_N, are *not* symmetric around $\bar{x}$.

On the right column of Figure 4 we show, for the same values of N, the distributions $\rho_{s_N}^{\mathrm{random}}$ of sums of N random values of x sampled from ρ_x. In contrast with the case of the Bernoulli map, $\rho_{s_N}^{\mathrm{random}}$ is here estimated numerically. As expected, the distributions of random sampling sums are now symmetric with respect to zero, and exhibit a fast convergence to the asymptotic Gaussian.

Figure 5 shows D_G and D_{random} for the fully chaotic logistic map, as functions of N. Since, as explained above, $\sigma_{s_N}^2$ equals σ_s^2 for all N, now D_{G_N} coincides with D_G. Due to the symmetry of ρ_x with respect to its mean value, the arguments given in Appendix B apply to this case, and D_{random} decays as N^{-2} for large N. The full straight line in the log–log plot of the figure has slope -2, confirming this prediction in the plotted range. Yet, the behavior of D_G is considerably different. It starts with a small increment between $N = 1$ and 2, where it attains a maximum, and thereafter decays rapidly up to $N \approx 20$. This decay corresponds to the interval of N for which the distribution ρ_{s_N} displays identifiable singularities. For $N \gtrsim 20$, the singularities start to overlap, and the distribution ρ_{s_N} varies more smoothly and displays a well-defined asymmetric bell-shaped profile. In this zone, the decay of D_G is slower and approximately behaves as N^{-1}, as illustrated by the dashed straight segment of slope -1. As shown in Appendix B, a decay as N^{-1} is expected for the KLD of the distribution of random sampling sums when the distribution of the individual summands is not symmetric with respect to the mean value. If the disparate dependence on N between D_G and D_{random} persists as N grows beyond the range considered here, their relative difference would increase indefinitely for $N \to \infty$.

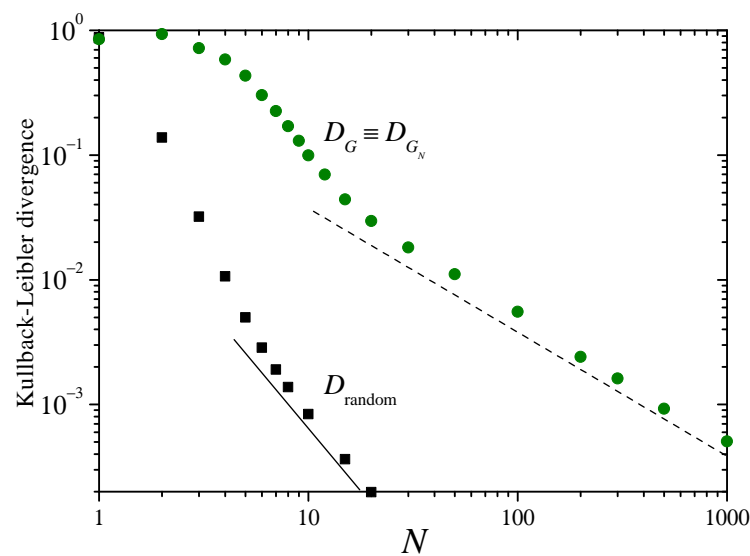**Figure 5.** The Kullback–Leibler divergence D_G for the logistic map of Equation (16) in the regime of full chaos, $\lambda = 4$, and D_{random}, as a functions of N. In this case, D_{G_N} coincides with D_G. The full and dashed straight lines have slopes -2 and -1, respectively.

Although still chaotic, other values of λ in Equation (16) give rise to qualitatively different dynamical features in the logistic map. For $\lambda = 3.828$, which is our next case of study, the dynamics are intermittent. Just above this value of λ, at $\lambda_3 = 1 + 2\sqrt{2} \approx 3.8284$, the logistic map enters the largest stability window within its chaotic regime, where $x(t)$ becomes asymptotically locked in a period-3 orbit. For $\lambda \lesssim \lambda_3$, the vicinity of the critical point manifests itself in the form of intermittent behavior for $x(t)$. Namely, the dynamics alternate intermittently between intervals of "turbulent" evolution, where its behavior is conspicuously chaotic, and "laminar" evolution, where $x(t)$ remains temporarily close to

the period-3 orbit, but eventually departs away from it. The left panel of Figure 6 shows 900 successive iterations of $x(t)$ for the above value of λ, illustrating both kinds of behavior.
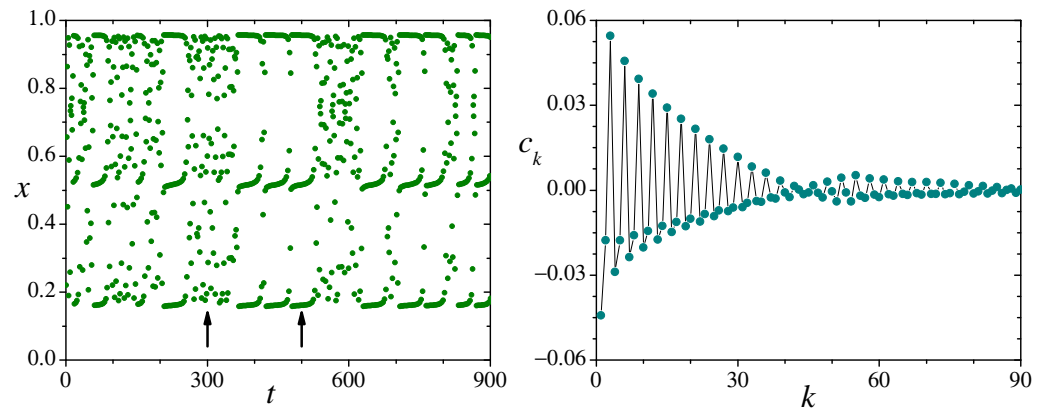

Figure 6. Left: 900 successive iterations of the logistic map, Equation (16), in the intermittent regime, $\lambda = 3.828$. The arrows at $t = 300$ and 500 point at "turbulent" and period-3 "laminar" intervals, respectively. **Right**: The correlation $c_k = \overline{[x(t) - \bar{x}][x(t + k) - \bar{x}]}$ as a function of k in the same intermittent regime, calculated numerically from sequences of 10^7 iterations of $x(t)$. Symbols are connected by lines to facilitate visualization.

For $\lambda = 3.828$, no analytical description of the logistic map exists, and we must resort to numerical techniques. As inferred from the left panel of Figure 6, in this case, $\rho_x(x)$ covers only a portion of the interval $[0, 1]$, between $x \approx 0.157$ and 0.957, and displays three peaks near the values of x in the period-3 orbit. Our numerical estimations for the mean value and the variance are $\bar{x} \approx 0.593$ and $\sigma_x^2 \approx 0.0864$. In principle, the variance of the sums s_N could be obtained from Equations (6) and (7) by numerically computing the correlations $c_k = \overline{[x(t) - \bar{x}][x(t + k) - \bar{x}]}$. These quantities, however, exhibit sharp oscillations and slow convergence as k grows, as well as persistent fluctuations for large k. The right panel in Figure 6 shows c_k up to $k = 90$. In practice, such features make impossible the evaluation of the variances $\sigma_{s_N}^2$ and σ_s^2 using the sums in Equations (6) and (7). We therefore resort to their direct numerical calculation using the values of $s_N(t)$ obtained from successive map iterations. In particular, our estimation for the variance of the sums in the limit $N \to \infty$ is $\sigma_s^2 \approx 0.0403$.

Colored symbols in the main panel of Figure 7 stand for D_G in the case of the logistic map with $\lambda = 3.828$, as a function of N. As with full chaos (cf. Figure 5), two distinct decay regimes are identifiable. Moreover, the behavior for $N \lesssim 50$ now contains signatures of the pseudo-periodic nature of the mapping in the "laminar" intervals, namely, the relatively large values of D_G when N is a multiple of 3 (triangles). In fact, for those values of N, the distributions ρ_{s_N} are narrower and sharper than for the remaining values, giving rise to higher KLDs. This is clearly illustrated by the dependence of the variance $\sigma_{s_N}^2$ on N, shown in the inset of the figure. After an abrupt initial decay, $\sigma_{s_N}^2$ displays oscillations of period 3, which progressively damp out as N grows. For $N \gtrsim 50$, the difference in D_G for multiples of 3 rapidly smooths out, as the KLD enters a regime where it decays approximately as N^{-1}, as indicated by the dashed segment of slope -1.

For this case of intermittent dynamics, we have also calculated D_{G_N}, finding qualitatively the same behavior as for D_G. As a matter of fact, D_G and D_{G_N} typically differ from each other in just about a 10%. Thus, for the sake of clarity, the numerical estimations of D_{G_N} were not included in Figure 7. As for the KLD of the distribution of random sampling sum, D_{ramdom}, the results of Appendix B indicate that it should decay as N^{-1} for large N. This behavior, however, has not yet been reached in the range of values displayed in Figure 7. Assuming nevertheless that this is the asymptotic dependence of D_{random}, our results suggest that the KLD for random sampling sums is no less than three orders of magnitude smaller than D_G for large N.
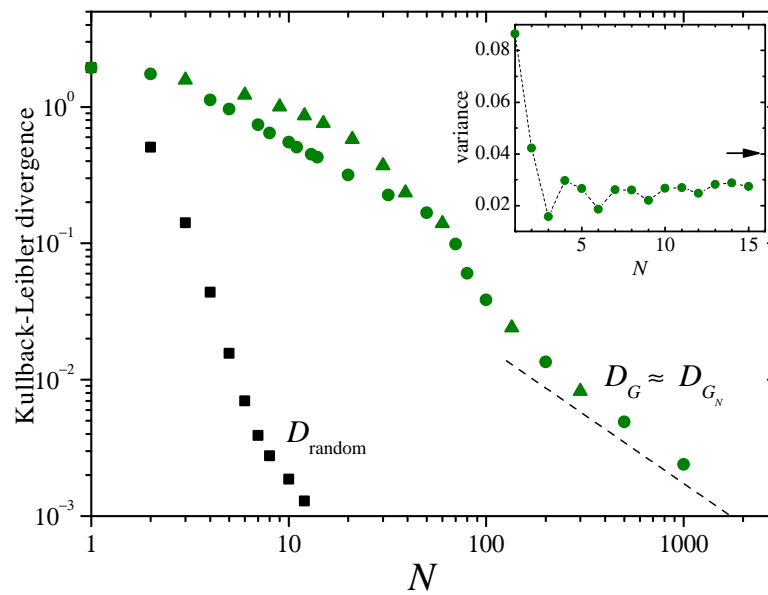
Figure 7. The Kullback–Leibler divergences D_G for the logistic map (16) in the intermittent regime, $\lambda = 3.828$, and D_{random}, as functions of N. For the former, triangles correspond to values of N which are multiples of 3. The slope of the dashed straight line is -1. Inset: Numerical results for the variance $\sigma_{s_N}^2$ of the sums s_N, as a function of N. The arrow to the right indicates the variance obtained for large N. Symbols are connected by dashed lines to facilitate visualization.

In summary, both for $\lambda = 4$ and 3.828, the main difference between the statistics of the sums s_N obtained from the iteration of the logistic map and from a random sampling of the corresponding invariant measures, as N grows, resides in their disparate rates of approach towards the asymptotic distribution. Within the range of N considered in our numerical calculations, the decay of D_G as N^{-1} can be qualitatively understood by the lack of symmetry in the invariant measures although, strictly speaking, the corresponding result in Appendix B holds for random sampling only.

Both when $\lambda = 4$ and 3.828, for $N \gtrsim 20$, the difference between D_G and D_{random} is well above two orders of magnitude. In the intermittent case, moreover, the pseudo-periodic character of the "laminar" dynamics reveals itself in the form of oscillations in D_G for small N, which are naturally absent in D_{random}. Plausibly, pseudo-periodicity is also responsible for the slow decrease in D_G during the oscillatory regime. Intermittency degrades the mixing properties of the mapping since, during the pseudo-periodic intervals, the dynamics only explore a reduced portion of the available range in x.

5. A Fat-Tailed Invariant Distribution

Much like the standard CLT, the CLT for deterministic systems can be generalized to the situation where the variance of the relevant variable x diverges [11]. In particular, this is the case of invariant distributions with a sufficiently slow algebraic decay for large $|x|$: $\rho_x(x) \sim |x|^{-\alpha-1}$ with $0 < \alpha < 2$. Under the same hypotheses of ergodicity and aperiodicity stated in Section 2, and assuming for simplicity that $\bar{x} = 0$—for instance, due to the symmetry of $\rho_x(x)$ around zero—the distribution of the sums

$$s_N(t) = \frac{1}{N^{1/\alpha}} \sum_{k=1}^{N} x(t+k-1) \tag{18}$$

[cf. Equation (3)] converges to a stable distribution given by the Fourier anti-transform of $Q_{\gamma_s}(k) = \exp(-\gamma_s^\alpha |k|^\alpha)$, for some value of the dispersion parameter γ_s. The result for distributions with finite variance is re-obtained in the limit $\alpha = 2$, with $\gamma_s \equiv \sigma_s$ as defined in Equation (7).

In this section, we give an example of convergence toward a stable distribution different from a Gaussian in the case of a map with a fat-tailed invariant distribution decaying as $|x|^{-2}$ for large $|x|$ (i.e., $\alpha = 1$). This specific case has the analytical advantage that the stable distribution predicted by the CLT can be explicitly written out, namely,

$$C_{\gamma_s}(s) = \frac{1}{\pi} \frac{\gamma_s}{\gamma_s^2 + s^2},\tag{19}$$

which is nothing but the Cauchy (or Lorentzian) distribution. Like the Gaussian, the Cauchy distribution is a maximum entropy distribution, but with a different constraint.

To obtain a deterministic chaotic map with a variable distributed following a fat-tailed function, we use the ad hoc procedure of applying a suitable transformation to a map whose invariant distribution is known in advance. Specifically, we take the Bernoulli map of Equation (10) with $m = 2$, for which we know that the invariant distribution is the function given by Equation (11), and introduce a change of variables that transforms this function into the desired fat-tailed profile. This is formally achieved by defining the two-variable map

$$\begin{cases} u(t+1) = \{2u(t)\}, \\ x(t+1) = \tau[\{2u(t)\}], \end{cases}\tag{20}$$

where

$$\tau(u) = \begin{cases} (2u-1)/2u & \text{for } 0 < u \leq 1/2, \\ (2u-1)/2(1-u) & \text{for } 1/2 \leq u < 1 \end{cases}\tag{21}$$

transforms a variable u with uniform distribution in $(0,1)$ into a variable $\tau \in (-\infty, \infty)$ with distribution $\rho_\tau(\tau) = 1/2(1 + |\tau|)^2$. By construction, thus, the invariant measure of variable x in map (20) is

$$\rho_x(x) = \frac{1}{2(1 + |x|)^2},\tag{22}$$

with x varying from $-\infty$ to ∞.

By analyzing the behavior of the Fourier transform of $\rho_x(x)$ near the origin, it is possible to obtain the dispersion parameter for the Cauchy distribution of sums of independently chosen values of x, which turns out to be $\gamma_s = \pi/2$. Unfortunately, the value of γ_s when the summands are successive iterations of x in map (20) cannot be found analytically in an explicit way. However, we have numerically found that, for $N \to \infty$, the dispersion parameter again coincides with $\gamma_s = \pi/2$ to a high precision. This is the value of γ_s that we use to compute the KLD $D_C = D(\rho_{s_N} || C_{\gamma_s})$ between the distribution of the sums s_N of Equation (18) and the Cauchy distribution (19). In addition, we do not have a practical procedure to assign a value to the dispersion parameter when the number of summands N is finite. Therefore, in the present case, we do not calculate a quantity analogous to the KLD D_{G_n} of Sections 3 and 4. Regarding D_{random}, due to the non-analytic behavior of the Fourier transform of $\rho_x(x)$ at the origin, it is now not possible to use the procedure of Appendix B to predict how this KLD decreases as N grows. Our analysis must thus rely on numerical results.

In Figure 8, we show the distributions $\rho_{s_N}(s_N)$ (left column) and $\rho_{s_N}^{\text{random}}(s_N)$ (right column) for three small values of N. Light-colored curves correspond to the expected asymptotic Cauchy distribution, given by Equation (19) with $\gamma_s = \pi/2$. Note that for $N = 2$, due to the peak at $s_N = 0$, the difference between $\rho_{s_N}^{\text{random}}$ and the asymptotic distribution seems to be larger than that of ρ_{s_N}. The KLDs, however, reveal that $\rho_{s_N}^{\text{random}}$ is slightly closer to the Cauchy distribution (see Figure 9). For $N = 10$, it is already clear that the approach to the Cauchy distribution is faster for the random sampling sums. Comparison with the results for the Bernoulli and the logistic maps (cf. Figures 1 and 4) suggest however that, in the present situation, the convergence to the corresponding asymptotic distribution is considerably slower than for those cases.
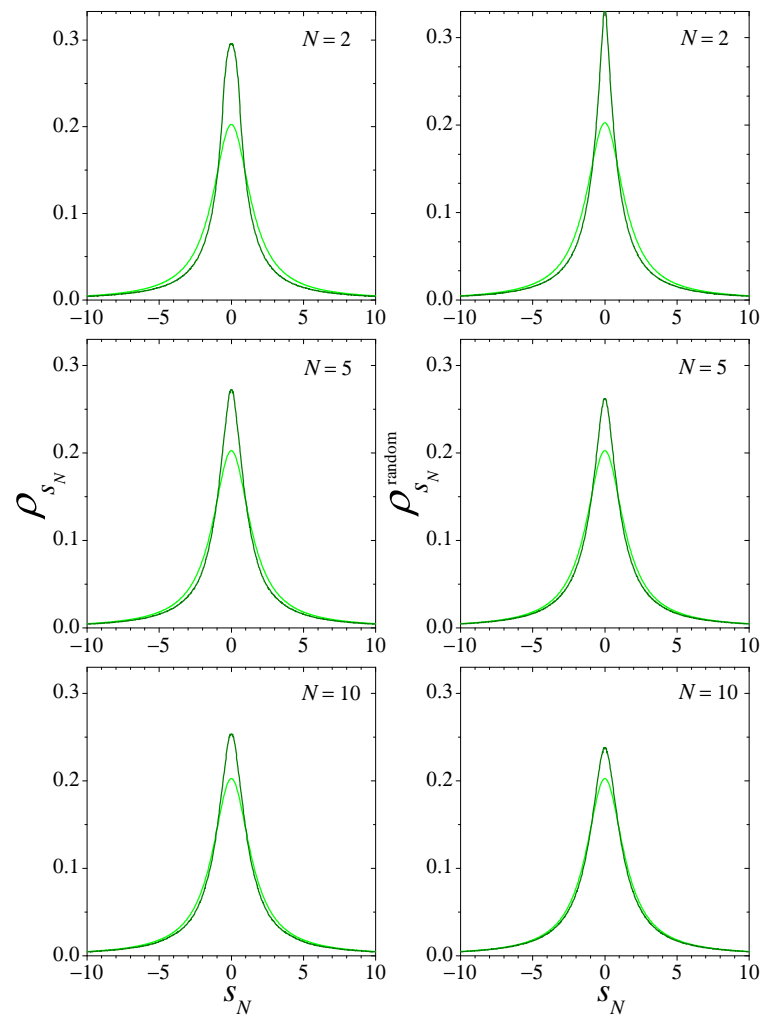
Figure 8. As in Figure 4, for the sums of Equation (18) with $x(t)$ obtained from map (20). Note that the scales are the same in all plots.

Figure 9 presents numerical results for the KLDs D_C and D_{random}. In order to have significant statistics in the construction of the 1000-column histogram that represents $\rho_{s_N}(s)$ from 10^7 samples of the sums s_N, we have cut off the interval of variation of s_N to $(-10, 10)$, disregarding samples outside that interval. Otherwise, for the fat-tailed distributions involved in the present case, the calculation of the KLDs would be dominated by sampling fluctuations for large values of $|s_N|$. Along most of the range of N spanned by the figure, both D_C and D_{random} exhibit rather well defined power–law decays. Their different exponents, however, make that they progressively diverge from each other as N grows, while D_{random} approximately decays as N^{-1}; as illustrated by the full straight line of slope -1, a linear fitting of D_C for $N \geq 2$, shown as a dashed line, points to a slower decay with a nontrivial exponent: $N^{-0.68}$. This result suggests that the convergence to an asymptotic distribution for the sums s_N in the case of fat-tailed invariant measures may generally be characterized by unusual exponents in the decay of the KLD. This conjecture will be thoroughly explored in future work, through both analytical and numerical means.
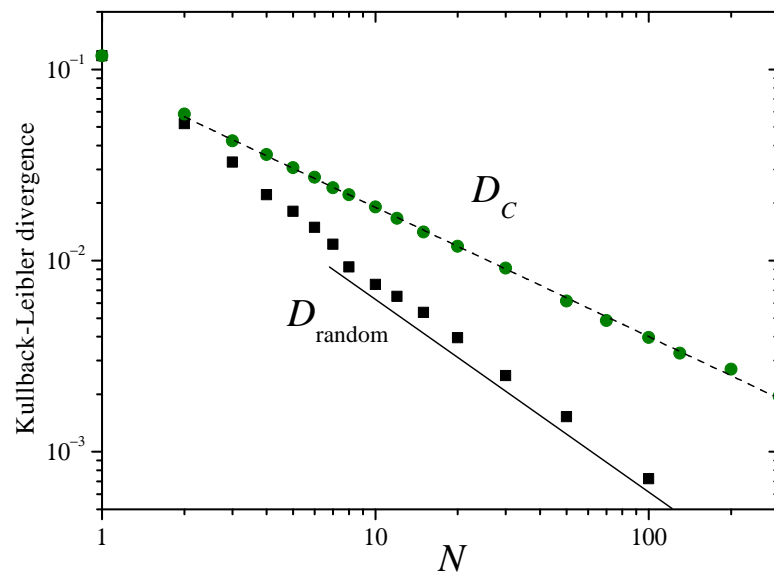
Figure 9. The Kullback–Leibler divergences D_C and D_{random} for the distributions of the sums of Equation (18), with the values of x obtained from the map (20) and the distribution of Equation (22), respectively. The full straight line has slope -1, and the dashed line, with slope -0.68, is a linear fitting of D_C for $N \geq 2$.

6. Conclusions

We here analyzed the convergence to the asymptotic probability density distribution ρ_s of a succession of distributions ρ_{s_N} for a conveniently scaled sum of N samples obtained from iterations of a deterministic map. Previous analytical studies had established that a modified version of the central limit theorem (CLT) exists for these cases. Yet, as far as we know, the convergence to the limit had not yet been characterized. Here, we studied several archetypal examples that expose a variety of ways the limiting distributions are approached.

Our characterization was based on the behavior of the Kullback–Leibler divergence (KLD) D_G between ρ_s and ρ_{s_N}, in that specific order. With this choice, the KLD equals the number of extra bits required to encode a sample from ρ_s if the code has been optimized for ρ_{s_N}. The CLT for sums of random samples with finite variance predicts a KLD D_{random} that decreases as N^{-2} if each sample is drawn from a distribution that is symmetric around its mean value, and as N^{-1} if it is not. This is a bold statement, since an infinitesimal modification may suffice to turn a symmetric distribution into an asymmetric one, so even a minute modification would suffice to change the entire asymptotic behavior of the KLD—the change, however, would only become relevant at increasingly larger values of N, as the asymmetry tended to disappear. We are not aware of an analogous theoretical prediction for the case of correlated samples, but the results presented have revealed similar behaviors: D_G decreased as N^{-2} for the Bernoulli map, for which the sums are distributed symmetrically around their mean value, and as N^{-1} for the logistic map, where the distributions are asymmetric.

In both the Bernoulli and the logistic map, the rates at which ρ_{s_N} approached the asymptotic distribution increased with the strength of mixing. Moreover, for the intermittent logistic map, where mixing is virtually absent during pseudo-periodic intervals, convergence to the asymptotic distribution was particularly slow. Therefore, even though all the explored examples were equally deterministic, their behavior differed considerably. Details in the chaotic dynamics are crucial to the behavior of D_G for large N.

The convergence of ρ_{s_N} in the Bernoulli map could be divided into two stages, one in which the distribution acquired an approximately Gaussian profile, and a subsequent one, in which the variance was adjusted to approach its asymptotic value. Remarkably, in the second stage and for sufficiently large N, the divergence $D_G(N)$ was equal to $D_{\text{random}}(\alpha N)$ with $\alpha \approx 3.74$, implying that each sample of the deterministic map was as informative

about the asymptotic distribution as α random samples. This equivalence could not be established in the other explored examples, since in all of them, D_{random} and D_G decreased with N with different power laws. No re-scaling procedure, hence, could transform one into the other.

The last example involved variables with divergent variance. In this case, the derivation of Appendix B is no longer valid, and no theoretical formulation describing how D_{random} tends to zero is known to us. Our numerical explorations revealed a behavior proportional to N^{-1} for D_{random}, even for samples drawn from distributions that are symmetric around their mean values. The deterministic counterpart D_G exhibited an even slower evolution, at a rate that is also slower than the one observed in the cases of finite variance.

In conclusion, in all the examples explored here, the asymptotic trend of the KLD behaved as a power law. Different deterministic maps yielded different exponents, displaying a variety of behaviors. The factors that influenced the exponents were (a) the strength of mixing in the chaotic map, (b) the tendency of the system to evolve near periodic orbits, and (c) the tails of the distribution of individual variables. We stress that the open question of establishing a quantitative connection between the rate of mixing, on one hand, and of KLD decay, on the other, remains an interesting problem for future work. Remarkably, except for the logistic map in the intermittent regime, all the maps explored here are related to each other by simple, nonlinear transformations. Despite these deterministic functional relations, their nonlinear nature determines differences in the statistical behavior of the sums of samples drawn from each map, with a large impact on the convergence towards their asymptotic distributions.

Author Contributions: Conceptualization, D.H.Z. and I.S.; Formal analysis, D.H.Z. and I.S.; Investigation, D.H.Z. and I.S.; Writing—review & editing, D.H.Z. and I.S. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Data are contained within the article.

Appendix A. Calculation of the Variances of N-Term Sums, Equation (6)

Appendix A.1. The Bernoulli Map

To calculate the variances $\sigma^2_{s^{(N)}}$ of the sums $s_t^{(N)}$ of Equation (3), as given by Equation (6), it is necessary to compute the correlations

$$c_k = \overline{[x(t) - \bar{x}][x(t+k) - \bar{x}]} = \int (x - \bar{x})\left[f^{(k)}(x) - \bar{x}\right]\rho_x(x)dx, \tag{A1}$$

where $f^{(k)}(x)$ is the k-th self-iteration of $f(x)$. For the Bernoulli map, Equation (10), $f^{(k)}(x)$ can be explicitly given as a piece-wise linear function over the interval $[0, 1)$:

$$f^{(k)}(x) = m^k\left[x - (j-1)m^{-k}\right], \quad \text{for } (j-1)m^{-k} \leq x < jm^{-k}, \tag{A2}$$

and $j = 1, 2, \ldots, m^k$. Taking into account that $\rho_x(x) = 1$ over the same interval, with $\bar{x} = 1/2$, the correlation in Equation (A1) turns out to be

$$c_k = \frac{1}{12}m^{-k}. \tag{A3}$$

Inserting this result in Equation (6), the variances of Equation (12) are straightforwardly obtained. Equation (A3) shows that, for the Bernoulli map, correlations between successive values of $x(t)$ decay exponentially with the span k, decreasing the faster the larger m is.

Appendix A.2. The Logistic Map in the Fully Chaotic Regime

The correlations $c_k = \overline{[x(t) - \bar{x}][x(t+k) - \bar{x}]}$ for the fully chaotic logistic map, $x(t+1) = 4x(t)[1 - x(t)]$, can be conveniently computed by exploiting the exact solution of the map [20,21],

$$x(t) = \sin^2(2^t \xi_0), \tag{A4}$$

where ξ_0 is determined using the initial condition $x(0)$ through the relation $x(0) = \sin^2 \xi_0$. Recalling that $\rho_x(x) = [\pi\sqrt{x(1-x)}]^{-1}$ and $\bar{x} = 1/2$, Equation (A1) implies

$$c_k = -\frac{1}{4} + \frac{1}{\pi} \int_0^1 \sqrt{\frac{x}{1-x}} \sin^2\left(2^k \arcsin\sqrt{x}\right) dx. \tag{A5}$$

The integral in this equation may look somehow intimidating but, using the change of variables $x = \sin^2 \xi$, it becomes the much simpler form $2\int_0^{\pi/2} \sin^2 \xi \sin^2(K\xi) d\xi$, with $K = 2^k$. Now, it can be easily shown—for instance, by induction over K—that the integral equals $\pi/8$ for all integers $K > 1$. From this result, it follows that $c_k = 0$ for all k. Remarkably, therefore, successive iterations of the logistic map in the fully chaotic regime are linearly uncorrelated with each other, although their functional correlation is obviously very large. Thus, the variance of the sums $s_t^{(N)}$ is

$$\sigma_{s_N}^2 = \sigma_x^2 = \frac{1}{8} \tag{A6}$$

for all N.

Appendix B. Kullback–Leibler Divergence for the Distribution of Random Sampling Sums

According to the Berry–Esséen theorem [22,23], the difference between the distribution for the sum of N independent random variables and the Gaussian predicted by the standard central limit theorem decays as $1/\sqrt{N}$ or faster as N grows. We show in this Appendix that, when the distribution of the individual random variables $\rho_x(x)$ admits a cumulant expansion—i.e., when the logarithm of its Fourier transform can be expanded in powers of its variable—that difference decays as $1/\sqrt{N}$ if $\rho_x(x)$ is asymmetric with respect to the mean value $\bar{x}$, and as $1/N$ if it is symmetric. This implies that the Kullback–Leibler divergence D_{random} defined in the main text decays as $1/N$ in the former case, and as $1/N^2$ in the latter. In the distributions considered in the main text, the symmetry with respect to the mean value is verified for the Bernoulli map and for the logistic map in the fully chaotic regime.

Without generality loss, we assume that the mean value over the distribution $\rho_x(x)$ of the individual random variables is zero. For the sums $s = \sum_{i=1}^N x_i / \sqrt{N}$, where x_i are independent samples of ρ_x, the distribution $\rho_s(s)$ results from the N-th order self-convolution of $\rho_x(x)$. This operation is most conveniently expressed in terms of the characteristic functions (Fourier transforms) $G_x(k)$ and $G_s(k)$ of, respectively, $\rho_s(s)$ and $\rho_x(x)$. Namely, $G_s(k) = [G_x(k/\sqrt{N})]^N$, or

$$\ln G_s(k) = N \ln G_x\left(\frac{k}{\sqrt{N}}\right) = N \sum_{j=1}^\infty \left(-\frac{ik}{\sqrt{N}}\right)^j \frac{\kappa_j}{j!}, \tag{A7}$$

where the sum in the right-hand side is the power expansion of $\ln G_x(k)$ around $k = 0$, which we assume to exist, and κ_j is the j-th order cumulant of $\rho_x(x)$ [24]. We recall that $\kappa_1 = \bar{x} = 0$, $\kappa_2 = \sigma_x^2$ is the variance of x over ρ_x, and $\kappa_3 = \overline{(x - \bar{x})^3}$.

Using this information, the anti-transform of $G_s(k)$ can be written as

$$
\begin{aligned}
\rho_s(s) &= \frac{1}{2\pi} \int_{-\infty}^\infty \exp\left(iks - \frac{\sigma_x^2}{2}k^2\right) \exp\left[\sum_{j=3}^\infty \frac{(-ik)^j}{N^{-1+j/2}} \frac{\kappa_j}{j!}\right] dk \\
&\equiv \frac{\exp(-s^2/2\sigma_x^2)}{\sqrt{2\pi\sigma_x^2}} + \Delta\rho_s(s),
\end{aligned} \tag{A8}
$$

with

$$\Delta\rho_s(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(iks - \frac{\sigma_x^2}{2}k^2\right) \left[\sum_{j=3}^{\infty} \frac{(-ik)^j}{N^{-1+j/2}} \frac{\kappa_j}{j!} + \cdots\right] dk, \tag{A9}$$

where the ellipsis stands for higher-order terms in the power expansion of the second exponential in the integrand of Equation (A8). Note that $\Delta\rho_s(s)$ is nothing but the difference between $\rho_s(s)$ and the asymptotic Gaussian distribution $G_{\sigma_x}(s)$ and that, due to normalization, it must verify $\int \Delta\rho_s(s)ds = 0$. If $\rho_x(x)$ is asymmetric around zero, the third-order cumulant κ_3 is different from zero, and the leading term in powers of N in $\Delta\rho_s(s)$ is given by the summand with $j = 3$ in Equation (A9), which implies $\Delta\rho_s \sim 1/\sqrt{N}$. If, on the other hand, $\rho_x(x)$ is symmetric around zero, we have $\kappa_3 = 0$. In this case, the sum effectively starts at $j \geq 4$, and $\Delta\rho_s$ decreases as $1/N$ or faster. Of course, if ρ_x is Gaussian from the start, all the higher-order cumulants vanish, and $\Delta\rho_s(s)$ is trivially equal to 0 for all N.

For the Kullback–Leibler divergence we have, from Equation (8),

$$\begin{aligned}
D_{\text{random}} &= D(\rho_s||G_{\sigma_x}) = \int [G_{\sigma_x}(s) + \Delta\rho_s(s)] \log_2\left[1 + \frac{\Delta\rho_s(s)}{G_{\sigma_x}(s)}\right] ds \\
&\approx \int [G_{\sigma_x}(s) + \Delta\rho_s(s)] \frac{\Delta\rho_s(s)}{G_{\sigma_x}(s)} ds = \int \frac{[\Delta\rho_s(s)]^2}{G_{\sigma_x}(s)} ds,
\end{aligned} \tag{A10}$$

where the approximation holds if $\Delta\rho_s(s)$ is sufficiently small. If the distribution $\rho_x(x)$ is asymmetric around zero, since $\Delta\rho_s(s)$ decays asymptotically as $1/\sqrt{N}$, then the decay of D_{random} turns out to be as $1/N$. If it is symmetric, D_{random} decays as $1/N^2$ or faster, depending on whether the subsequent cumulants vanish or not.

References

1. Lévy, P. *Theorie de L'addition des Variables Aleatoires*; Gauthier-Villars: Paris, France, 1937.
2. Gnedenko, B.V.; Kolmogorov, A.N. *Limit Distributions for Sums of Independent Random Variables*; Addison-Wesley: Cambridge, MA, USA, 1954.
3. Uchaikin, V.V.; Zolotarev, V.M. *Chance and Stability: Stable Distributions and their Applications*; Walter de Gruyter: Boston, MA, USA, 2011.
4. Araujo, A.; Giné, E. *The Central Limit Theorem for Real and Banach Valued Random Variables*; Wiley: New York, NY, USA, 1980.
5. Billingsley, P. *Probability and Measure*; Wiley: New York, NY, USA, 1995.
6. Burton, R.; Denker, M. On the central limit theorem for dynamical systems. *Trans. Am. Math. Soc.* **1987**, *302*, 715. [CrossRef]
7. Denker, M. The central limit theorem for dynamical systems. *Banach Cent. Publ.* **1989**, *1*, 33. [CrossRef]
8. Liverani, C. Central limit theorem for deterministic systems. *Pitman Res. Notes Math. Ser.* **1996**, *362*, 56.
9. Nicol, M.; Török, A.; Vaienti, S. Central limit theorems for sequential and random intermittent dynamical systems. *Ergod. Theor. Dyn. Syst.* **2018**, *38*, 1127. [CrossRef]
10. Kosloff, Z.; Volný, D. Local limit theorem in deterministic systems. *Ann. Inst. Henri Poincaré Probab. Stat.* **2022**, *58*, 548. [CrossRef]
11. Kosloff, Z.; Volný, D. Stable CLT for deterministic systems. *arXiv* **2023**, arXiv:2211.03448.
12. Buzzi, J. Chaos and Ergodic Theory. In *Encyclopedia of Complexity and Systems Science*; Meyers, R., Ed.; Springer: New York, NY, USA, 2009.
13. Wouters, J. Deviations from Gaussianity in deterministic discrete time dynamical systems. *Chaos* **2020**, *30*, 023117. [CrossRef] [PubMed]
14. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; Wiley-Interscience: Hoboken, NJ, USA, 2006.
15. Irwin, J. O. On the frequency distribution of the means of samples from a population having any law of frequency with finite moments, with special reference to Pearson's type II. *Biometrika* **1927**, *19*, 225. [CrossRef]
16. Hall, P. The distribution of means for samples of size N drawn from a population in which the variate takes values between 0 and 1, all such values being equally probable. *Biometrika* **1927**, *19*, 240. [CrossRef]
17. May, R.M. Simple mathematical models with very complicated dynamics. *Nature* **1976**, *261*, 459. [CrossRef] [PubMed]
18. Tsuchiya, T.; Yamagishi, D. The complete bifurcation diagram for the logistic map. *Z. Naturforsch.* **1997**, *52*, 513. [CrossRef]
19. Jakobson, M. Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. *Comm. Math. Phys.* **1981**, *81*, 39. [CrossRef]
20. Schroder, E. Über iterirte Functionen. *Math. Ann.* **1870**, *3*, 296. [CrossRef]
21. Maritz, M.F. A note on exact solutions of the logistic map. *Chaos* **2020**, *30*, 033136. [CrossRef] [PubMed]
22. Hazewinkel, M. (Ed.) *Encyclopaedia of Mathematics*; Reidel: Dordrecht, The Netherlands, 1988.

23. Korolev, V.Y.; Shevtsova, I.G. On the upper bound for the absolute constant in the Berry–Esseen inequality. *Theor. Prob. Appl.* **2010**, *54*, 638. [CrossRef]
24. Gardiner, C.W. *Handbook of Stochastic Methods*; Springer: Berlin/Heidelberg, Germany, 1997.