Asymptotic scaling behavior of block entropies for an intermittent process

Jan Freund*

Institute of Physics, Humboldt-University Berlin, Invalidenstraβe 110, D-10115 Berlin, Germany

(Received 20 July 1995)

Intermittent systems play a prominent role in the field of dynamical phase transitions. Their extraordinary characteristics also show up when applying the concept of block entropies to symbol sequences which are generated through the method of symbolic dynamics. We investigate the dependence of those dynamical entropies on the block length n. In the asymptotic limit, i.e., for $n\to\infty$, an analytic treatment is possible when starting from the assumption of independent laminar laps. The results are important for a refined characterization of intermittent systems. Moreover, they bring intermittent systems in contact with information-carrying sequences which exhibit a very special scaling behavior of dynamical entropies. [S1063-651X(96)10605-X]

PACS number(s): 02.50−r, 05.45.+b, 47.27.Cn, 47.27.Eq

I. INTRODUCTION

The notion of intermittency was first introduced in connection with hydrodynamic systems at the transition point from the laminar to turbulent regime. Manneville and Pomeau numerically solved the differential equations of the Lorenz model [1]. Above a threshold value for an external control parameter r they observed the transition from a stable periodic motion to a pattern where the *laminar* regime was interrupted by *chaotic* (or *turbulent*) *bursts*. An appropriate Poincaré plot revealed the underlying mechanism: the transition could be related to a fixed point losing its stability when the control parameter r exceeds the critical value. Of special importance was their observation that this scenario offered a *universal route to chaos* [2]. Additionally it provides a universal mechanism for $1/f$ noise in nonlinear systems [3]. Subsequently intermittency has been detected in many physical systems (see [4,5] and references therein).

According to the different ways a fixed point can lose its stability [4], three types of intermittency (I,II,III) are distinguished. Here, we will be mainly concerned with intermittency type I; in this case there exists a maximal length of laminar phases which will be denoted by m. The decisive properties of an intermittent system are the length distributions for the laminar (turbulent) lengths. It is a common assumption that the successive laps are independent of each other [4,6,7]; this assumption often is formulated by describing the dynamics as a *regenerative process*.

Whereas the length distribution of laminar phases sensitively depends on slight changes of the dynamics, the typical length of chaotic bursts does not. This is easy to understand since closing the narrow channel [4] a little further strongly affects the laminar creep but the chaotic behavior outside the channel region essentially remains the same. A proper rescaling of time (by the inverse average chaotic length) allows for an idealized description of the chaotic phases as a one step process.

The model we consider here relates to a time discrete intermittent dynamics. Successive laminar lengths i (for $i=1,\ldots,m$) are chosen independently according to a given distribution $p(i)$. In the context of symbolic dynamics [8] we only distinguish between the laminar region, symbol "0," and the chaotic region, symbol "1." Hence a typical symbol sequence will look like "0000000000100000010000000000001010001 . . . ." We regard the symbols "0" and "1" as elements of a *binary alphabet* $A=\{0,1\}$. In the general case we deal with an alphabet consisting of λ letters, i.e., $A=\{a_1,\ldots,a_\lambda\}$.

The statistics of subsequences $(c_1,\ldots,c_n)$ (with $c_i\in A$ $\forall i=1,\ldots,n$) of length n, also named *n words*, contains information about the system under consideration. The statistics can be derived either by analyzing the frequencies of n words excerpted from a sample string or by applying analytical methods. Here, we will use the last mentioned approach and hence do not have to care about finite sample size effects [9,10]. We denote the probability of an arbitrary n word by the symbol $p(c_1,\ldots,c_n)$.

The informational analysis of a dynamical system proceeds by inserting those n-word probabilities into functionals which fulfill the axioms of an information measure [11]. The measures we use in this work are the *block entropies* H_n ($n=1,2,\ldots$) [12] related to the Shannon information [13]:

$$H_n:=-\sum_{(c_1,\ldots,c_n)\in A^n} p(c_1,\ldots,c_n)\log_\lambda p(c_1,\ldots,c_n). \tag{1}$$

Choosing logarithms to base λ is favorable since then the inequality $0\leq H_n\leq n$ holds.

The H_n can be interpreted as the average uncertainty when trying to predict an n word (a trajectory segment of length $n\tau$). Dividing H_n by n yields the average uncertainty per symbol, i.e., $H(n):=H_n/n$. Its limiting value for $n\to\infty$ was named the *entropy of the source* [14]:

$$h:=\lim_{n\to\infty} H(n). \tag{2}$$

The entropy of the source is closely connected with the *Kolmogorov-Sinai entropy* [15,16].

*Electronic address: janf@summa.physik.hu-berlin.de

The notion of a *memory* assigned to a symbol sequence is linked with the improvement of the prediction of a symbol c_{n+1} given knowledge of its prehistory $(c_1,\ldots,c_n)$. This idea naturally involves conditional probabilities and, hence, the corresponding informational quantities h_n are named *conditional entropies*; they are defined by

$$h_n := \left\langle -\sum_{a_{n+1}\in A} p(c_{n+1}|c_1,\ldots,c_n) \times \log_\lambda p(c_{n+1}|c_1,\ldots,c_n) \right\rangle_{(c_1,\ldots,c_n)} \tag{3}$$

$$= H_{n+1} - H_n (n=1,2,\ldots). \tag{4}$$

Correlations existing between a subsequence $(c_1,\ldots,c_n)$ and the symbol c_{n+1} improve the chance of guessing c_{n+1} successfully, i.e., the average uncertainty of prediction is diminished. Thus the h_n will monotonically decline when increasing the length n of noticed prehistory $(c_1,\ldots,c_n)$. It can be shown that for *stationary and ergodic sources* [11,14] the limit of the h_n for $n\to\infty$ exists and coincides with the entropy of the source, i.e.,

$$\lim_{n\to\infty} h_n = h. \tag{5}$$

The profile of h_n is directly related to the *effective measure complexity* (EMC) defined by Grassberger [17]:

$$\mathrm{EMC} := \sum_{n=0}^{\infty} (h_n - h). \tag{6}$$

The specific way the h_n approach this limit can be used to characterize a dynamical system [18]. For *Markov sources of order m* [19] the conditional entropies already reach their limit for $n=m$, i.e., $h_n = h = H_{m+1} - H_m$ for all $n \geqslant m$ [19]. Hence such a Markov source possesses a memory of m steps [20].

For most systems the h_n decay exponentially [17,18,21], i.e., $(h_n - h) \sim \exp(-\gamma n)$. Then the inverse decay rate γ^{-1} can be defined as an effective memory; a Markovian approximation of sufficiently high order will be a very effective description.

A subexponential decay of the conditional entropies is an exceptional but rather interesting case. Szépfalusy and coworkers have shown that intermittent processes belong to this type of system [18,22]. Moreover, a subexponential decay was found for sequences which are not directly related to a dynamical system but instead are the result of an evolutionary process, e.g., literary texts and coded music [12,23–25]. This observation inspires the intriguing task of searching for a connection between intermittent dynamics and mechanisms underlying the creation of *natural symbol sequences*.

In this paper we report on the informational analysis of a binary sequence which was constructed by the aforementioned independent concatenation of laminar laps with randomly distributed lengths. As explained above, it mimics the symbolic dynamics of an intermittent system (type I). We present the result of a numerical simulation and, moreover, an analytic formula for the H_n valid in the asymptotic limit, i.e., for $n\to\infty$. The basic idea of this formula rests on the fact that the distribution of laminar lengths will be reflected by relative frequencies of subwords contained in a subsequence of length $n \gg \langle l \rangle$.

II. DERIVATION OF THE ASYMPTOTIC SCALING BEHAVIOR

In a previously published paper [26] we have already derived the expression to be evaluated yielding the block entropy H_n. Here just repeat the basic arguments.

A sample string is constructed as explained in the Introduction with ''0'' denoting the laminar phase and ''1'' the chaotic burst (the end of the laminar phase) (termination symbol). Hence an arbitrary n word excerpted from the sample string is a binary sequence. It can be decomposed into a collection of subwords of different lengths. Denoting the length of the ith subword within the n word by l_i we can symbolically represent this n word by

$$\underline{n} = (l_1,\ldots,l_k) \tag{7}$$

with the restriction

$$\sum_{i=1}^{k} l_i = n. \tag{8}$$

Special attention has to be paid to this restriction because the first and last subword within the n word are mostly cut out of a longer subwords found in the sample string. The probability of hitting a word of length i in the sample string is p_i and there are exactly i different letters which can be chosen as the starting position of the reading frame. Then follows the *bulk* of phases in the center of the n word and finally the right edge is determined by the restriction (8). It is important to notice that in the limit of sufficiently long n words the bulk will dominate the statistics. The different possibilities of placing the *intron* symbol yield a factor $\sum_{i=1}^{m} i p_i$ which is identified as the mean length and, henceforth, will be denoted by the symbol $\langle l \rangle$. Note that there is a maximum length denoted by m which means we are restricted to intermittency of type I.

Because of selecting the laminar phases independently of each other the probability for an arbitrary n word, denoted $P(\underline{n})$, can be written as

$$P(\underline{n}) \propto \prod_{i=1}^{k} p(l_i). \tag{9}$$

Since permuting the order of the laminar phases leaves this probability invariant there is a whole *class of equivalent words*. This fact motivates an occupancy number representation designating the whole class by the symbol $\underline{k} := (k_1,\ldots,k_m)$. Here $k_i \in \{0,1,\ldots,n\}$ stands for the number of laminar phases within a representative n word which have the length i. The restriction to a fixed word length (8) transforms in this representation to

$$\sum_{i=1}^{m} i k_i = n. \tag{10}$$

We can readily write down the contribution of a certain class to the n-word entropy:

$$\mathscr{H}(k_1,\ldots,k_m) := -\langle l\rangle \frac{(\Sigma_{i=1}^m k_i)!}{\Pi_{j=1}^m k_j!} \times \left\{\prod_{i=1}^m p(i)^{k_i}\right\} \log_2 \left\{\prod_{j=1}^m p(j)^{k_j}\right\}. \tag{11}$$

The first factor is linked to the starting phase l_1 as explained above. The next is a combinatorial expression reflecting the number of the members of the class $(k_1,\ldots,k_m)$. The third term is the probability of a representative of that class and the last is the usual information gained when noticing this n word.

Defining the class probability $\mathscr{P}(k_1,\ldots,k_m)$ by

$$\mathscr{P}(k_1,\ldots,k_m) := \langle l\rangle \frac{(\Sigma_{i=1}^m k_i)!}{\Pi_{j=1}^m k_j!}\left\{\prod_{i=1}^m p(i)^{k_i}\right\}, \tag{12}$$

Eq. (11) can equivalently be written as

$$\mathscr{H}(k_1,\ldots,k_m) = -\mathscr{P}(k_1,\ldots,k_m)\times\left(\sum_{i=1}^m k_i \log_2 p(i)\right). \tag{13}$$

Summing over all classes compatible with restriction (10) the complete n-word entropy yields

$$H_n = -\sum_{i=1}^m \langle k_i\rangle^{\dagger}_{(k_1,\ldots,k_m)} \log_2 p(i), \tag{14}$$

where we have used the abbreviation

$$\langle k_i\rangle^{\dagger}_{(k_1,\ldots,k_m)} := \sum_{k_1,\ldots,k_m}^{\dagger} \mathscr{P}(k_1,\ldots,k_m) k_i \tag{15}$$

and where the dagger is a reminder of the condition (10).

We account for the restriction (10) by inserting a Kronecker δ and additionally introduce a new variable N together with another Kronecker δ which imposes the constraint $N = \Sigma_{i=1}^m k_i$. Hence,

$$\langle k_i\rangle^{\dagger}_{(k_1,\ldots,k_m)} = \langle l\rangle \sum_{N=1}^n \sum_{k_1=0}^n \cdots \sum_{k_m=0}^n \times \frac{N!}{k_1!\cdots k_m!} p(1)^{k_1}\cdots p(m)^{k_m} \times \delta\left(\sum_{i=1}^m i k_i, n\right)\delta\left(\sum_{i=1}^m k_i, N\right) k_i. \tag{16}$$

The replacement

$$k_i = \imath \frac{\partial}{\partial\omega_i}[\exp(-i\underline{\omega}\cdot\underline{k})]|_{\underline{\omega}=\underline{0}} \tag{17}$$

means that we solve the problem using the characteristic function. [Throughout the text we adopt the following notation: $\imath$ is the imaginary unit, not to be confused with the index i; $\underline{a} := (a_1,\ldots,a_m)$ and $\underline{a}\cdot\underline{b} := \Sigma_{i=1}^m a_i b_i$]. This yields

$$\langle k_i\rangle^{\dagger}_{(k_1,\ldots,k_m)} = \imath\langle l\rangle \frac{\partial}{\partial\omega_i}\left(\sum_{N=1}^n \sum_{k_1=0}^n \cdots \sum_{k_m=0}^n \times \frac{N!}{k_1!\cdots k_m!} p(1)^{k_1}\cdots p(m)^{k_m} \times \delta\left(\sum_{i=1}^m i k_i, n\right)\delta\left(\sum_{i=1}^m k_i, N\right) \times \exp(-\imath\underline{\omega}\cdot\underline{k})\right)\Bigg|_{\underline{\omega}=\underline{0}}. \tag{18}$$

Defining now

$$F(\underline{\omega}) := \sum_{N=1}^n \sum_{k_1=0}^n \cdots \sum_{k_m=0}^n \frac{N!}{k_1!\cdots k_m!} p(1)^{k_1}\cdots p(m)^{k_m} \times \delta\left(\sum_{i=1}^m i k_i, n\right)\delta\left(\sum_{i=1}^m k_i, N\right)\exp(-\imath\underline{\omega}\cdot\underline{k}), \tag{19}$$

we can write Eq. (18) as

$$\langle k_i\rangle^{\dagger}_{(k_1,\ldots,k_m)} = \imath\langle l\rangle \frac{\partial}{\partial\omega_i}[F(\underline{\omega})]|_{\underline{\omega}=\underline{0}} \tag{20}$$

and Eq. (14) as

$$H_n = -\sum_{i=1}^m \imath\langle l\rangle \frac{\partial}{\partial\omega_i}[F(\underline{\omega})]|_{\underline{\omega}=\underline{0}} \log_2 p(i). \tag{21}$$

Hence the remaining task is to calculate the characteristic function $F(\underline{\omega})$.

The first step in this direction is to replace the sum by an integral, the discrete arguments $k_1,\ldots,k_m$ and N by real numbers, and accordingly the Kronecker δ's by δ distributions; i.e.,

$$F(\underline{\omega}) = \int_1^n dN \int_0^n dk_1 \cdots \int_0^n dk_m \frac{N!}{k_1!\cdots k_m!} \times p(1)^{k_1}\cdots p(m)^{k_m} \times \delta\left(\sum_{i=1}^m i k_i - n\right)\delta\left(\sum_{i=1}^m k_i - N\right)\exp(-\imath\underline{\omega}\cdot\underline{k}). \tag{22}$$

Next we use Stirling's formula for the factorials $x! = (x/e)^x\sqrt{2\pi x}\exp(\Theta_x/12x)$ with $0<\Theta_x<1$ and approximate $\Theta_x/12x \approx 0$. This is allowed for $n\to\infty$ since then (with very high probability) all the k_i for $i=1,\ldots,m$ and N eventually become sufficiently large. The precise mathematical argument can be found in [27]. This yields for the characteristic function

$$F(\underline{\omega})\approx\int_1^n dN\int_0^n dk_1\cdots\int_0^n dk_m \times\frac{\sqrt{2\pi N}}{\sqrt{2\pi k_1}\cdots\sqrt{2\pi k_m}}\exp\left(\sum_{i=1}^m k_i\ln\frac{p(i)}{k_i/N}\right) \times\delta\left(\sum_{i=1}^m ik_i-n\right)\delta\left(\sum_{i=1}^m k_i-N\right)\exp(-\imath\underline{\omega}\cdot\underline{k}). \tag{23}$$

For reasons of convenience we define $p_i:=p(i)$. Since in the asymptotic regime $n\to\infty$ the relative frequencies of the different laminar phases approach the underlying probabilities of the length distribution, we expand k_i/N around p_i, which corresponds to the substitution

$$k_i=N(p_i+\eta_i), \tag{24}$$

$$d^m\underline{k}dN=N^m d^m\underline{\eta}dN. \tag{25}$$

The η_i are the relative frequencies' deviations from the probabilities. We obtain

$$F(\underline{\omega})\approx\int_1^n dN\int_{-p_1}^{n/N-p_1}d\eta_1\cdots\int_{-p_m}^{n/N-p_m}d\eta_m\frac{N^m\sqrt{2\pi N}}{\sqrt{2\pi Np_1(1+\eta_1/p_1)}\cdots\sqrt{2\pi Np_m(1+\eta_m/p_m)}}\exp\left(-\sum_{i=1}^m N(p_i+\eta_i)\right. \times\left.\left[\frac{\eta_i}{p_i}-\frac{1}{2}\frac{\eta_i^2}{p_i^2}+\frac{1}{3}\frac{\eta_i^3}{p_i^3}-\cdots\right]\right)\delta\left(N\sum_{i=1}^m i\eta_i+N\langle l\rangle-n\right)\delta\left(N\sum_{i=1}^m\eta_i\right)\exp[-\imath N\underline{\omega}\cdot(\underline{p}+\underline{\eta})], \tag{26}$$

where the (infinite) series in square brackets corresponds to $\ln(1+\eta_i/p_i)$.

Now we choose the Fourier representation for both δ distributions:

$$\delta\left(N\sum_{i=1}^m i\eta_i+N\langle l\rangle-n\right) =\int_{-\infty}^{\infty}\frac{d\alpha}{2\pi}\exp\left[-\imath\alpha\left(N\sum_{i=1}^m i\eta_i+N\langle l\rangle-n\right)\right], \tag{27}$$

$$\delta\left(N\sum_{i=1}^m\eta_i\right)=\int_{-\infty}^{\infty}\frac{d\beta}{2\pi}\exp\left(-\imath\beta N\sum_{i=1}^m\eta_i\right). \tag{28}$$

In the limit $n\to\infty$ we retain just the most dominant terms in η, which corresponds to retaining only the quadratic term in the exponential (originating from the polynomial distribution; the linear term is absent because of the condition $\Sigma_{i=1}^m\eta_i=0$) and, furthermore, to approximating the brackets under the square roots by the factor 1. Of course, the δ distributions are not affected by any approximation. This yields

$$F(\underline{\omega})\approx\int_1^n dN\int_{-\infty}^{\infty}\frac{d\alpha}{2\pi}\int_{-\infty}^{\infty}\frac{d\beta}{2\pi}\int_{-p_1}^{n/N-p_1}d\eta_1\cdots \times\int_{-p_m}^{n/N-p_m}d\eta_m\frac{\sqrt{2\pi N}}{\sqrt{2\pi p_1/N}\cdots\sqrt{2\pi p_m/N}} \times\exp\left(-\frac{1}{2}\sum_{i=1}^m\frac{\eta_i^2}{p_i/N}\right) \times\exp\left[-\imath\alpha\left(N\sum_{i=1}^m i\eta_i+N\langle l\rangle-n\right)\right] \times\exp\left(-\imath\beta N\sum_{i=1}^m\eta_i\right)\exp[-\imath N\underline{\omega}\cdot(\underline{p}+\underline{\eta})]. \tag{29}$$

We see that the integrand is a product of m Gaussian functions, originating from the polynomial distribution, multiplied by $(3\times m)$ exponentials linear in the η_i, two originating from the δ distributions and one stemming from the characteristic function.

In the limit $n\to\infty$ the value of N tends to infinity too (in fact, in this case $N\to n/\langle l\rangle$). Since the variance of the Gaussians is given by $\sqrt{\sigma_i}=\sqrt{p_i/N}\sim1/\sqrt{N}$ the integrand eventually concentrates around values $\eta_i\approx0$, and a replacement of the limits $-p_i$ $(n/N-p_i)$ by $-\infty$ $(+\infty)$ results in negligible corrections. Performing the m integrations with respect to the η_i yields

$$F(\underline{\omega})\approx\int_1^n dN\int_{-\infty}^{\infty}\frac{d\alpha}{2\pi}\int_{-\infty}^{\infty}\frac{d\beta}{2\pi}\sqrt{2\pi N}\exp[-\imath N\underline{\omega}\cdot\underline{p}-\imath\alpha(N\langle l\rangle-n)]\exp\left(-\frac{1}{2}N\sum_{i=1}^{m}p_i(i\alpha+\beta+\omega_i)^2\right). \tag{30}$$

Next we perform the integration with respect to β. Defining

$$\sigma_{ll}:=\langle l^2\rangle-\langle l\rangle^2=\sum_{i=1}^{m}i^2p_i-\sum_{i,j=1}^{m}ijp_ip_j,$$

$$\sigma_{l\omega}:=\langle l\omega\rangle-\langle l\rangle\langle\omega\rangle=\sum_{i=1}^{m}i\omega_ip_i-\sum_{i,j=1}^{m}i\omega_jp_ip_j,$$

$$\sigma_{\omega\omega}:=\langle\omega^2\rangle-\langle\omega\rangle^2=\sum_{i=1}^{m}\omega_i^2p_i-\sum_{i,j=1}^{m}\omega_i\omega_jp_ip_j,$$

we obtain

$$F(\underline{\omega})\approx\int_1^n dN\int_{-\infty}^{\infty}\frac{d\alpha}{2\pi}\exp[-\imath N\underline{\omega}\cdot\underline{p}-\imath\alpha(N\langle l\rangle-n)]\times\exp\left(-\frac{1}{2}N(\alpha^2\sigma_{ll}+2\alpha\sigma_{l\omega}+\sigma_{\omega\omega})\right). \tag{31}$$

The integration with respect to α can be performed and we obtain

$$F(\underline{\omega})\approx\frac{1}{\sqrt{2\pi\sigma_{ll}}}\int_1^n\frac{dN}{\sqrt{N}}\exp(-\imath N\underline{\omega}\cdot\underline{p})\exp\left(-\frac{(\langle l\rangle-n/N)^2+\sigma_{ll}\sigma_{\omega\omega}-\sigma_{l\omega}^2-\imath\,2\sigma_{l\omega}(\langle l\rangle-n/N)}{2\sigma_{ll}/N}\right). \tag{32}$$

Performing the substitution $y:=\langle l\rangle-n/N$ yields

$$F(\underline{\omega})\approx\frac{1}{\langle l\rangle}\frac{1}{\sqrt{2\pi\sigma_{ll}\langle l\rangle/n}}\int_{\langle l\rangle-n}^{\langle l\rangle-1}\frac{dy}{(1-y/\langle l\rangle)^{3/2}}\exp\left(-\frac{(y^2+\sigma_{ll}\sigma_{\omega\omega}-\sigma_{l\omega}^2)+\imath(2\sigma_{ll}\underline{\omega}\cdot\underline{p}-2\sigma_{l\omega}y)}{(2\sigma_{ll}\langle l\rangle/n)(1-y/\langle l\rangle)}\right). \tag{33}$$

For the partial derivative with respect to ω_i we achieve

$$\imath\langle l\rangle\frac{\partial}{\partial\omega_i}[F(\underline{\omega})]|_{\underline{\omega}=\underline{0}}=\frac{n}{\langle l\rangle}p_i\times\frac{1}{\sqrt{2\pi\sigma_{ll}\langle l\rangle/n}}\int_{\langle l\rangle-n}^{\langle l\rangle-1}\frac{dy}{(1-y/\langle l\rangle)^{5/2}}\exp\left(-\frac{y^2}{(2\sigma_{ll}\langle l\rangle/n)}\frac{1}{(1-y/\langle l\rangle)}\right)\left(1+\frac{(\langle l\rangle-i)y}{\sigma_{ll}}\right). \tag{34}$$

In accordance with Eq. (21) we finally end up with

$$H_n=\frac{n}{\langle l\rangle}\left(-\sum_{i=1}^{m}p_i\log_2p_i\right)\mathcal{T}_0(n,\langle l\rangle,\sigma_{ll})+\frac{n}{\langle l\rangle}\left(-\sum_{i=1}^{m}p_i\log_2p_i\frac{\langle l\rangle-i}{\sigma_{ll}}\right)\mathcal{T}_1(n,\langle l\rangle,\sigma_{ll}) \tag{35}$$

$$=\frac{n}{\langle l\rangle}H[\mathbf{p}]\mathcal{T}_0(n,\langle l\rangle,\sigma_{ll})+\frac{n}{\langle l\rangle}A[\mathbf{p}]\mathcal{T}_1(n,\langle l\rangle,\sigma_{ll}), \tag{36}$$

where the functions $\mathcal{T}_0$ and $\mathcal{T}_1$ are defined by

$$\mathcal{T}_0(n,\langle l\rangle,\sigma_{ll}):=\frac{1}{\sqrt{2\pi\sigma_{ll}\langle l\rangle/n}}\int_{\langle l\rangle-n}^{\langle l\rangle-1}\frac{dy}{(1-y/\langle l\rangle)^{5/2}}\times\exp\left(-\frac{y^2}{(2\sigma_{ll}\langle l\rangle/n)}\frac{1}{(1-y/\langle l\rangle)}\right), \tag{37}$$

$$\mathcal{T}_1(n,\langle l\rangle,\sigma):=\frac{1}{\sqrt{2\pi\sigma_{ll}\langle l\rangle/n}}\int_{\langle l\rangle-n}^{\langle l\rangle-1}\frac{y\,dy}{(1-y/\langle l\rangle)^{5/2}}\times\exp\left(-\frac{y^2}{(2\sigma_{ll}\langle l\rangle/n)}\frac{1}{(1-y/\langle l\rangle)}\right). \tag{38}$$

Note that $H[\mathbf{p}]$ is just the entropy of the length distribution of the laminar phases. It can be shown that the functional $A[\mathbf{p}]$ vanishes for distributions that are symmetric with respect to the mean length $\langle l\rangle$; hence it can be regarded as an *asymmetry term*.

For sufficiently large n the integrands of $\mathcal{T}_0$ and $\mathcal{T}_1$ essentially are dominated by the Gaussian part. In the asymptotic limit $n\to\infty$ the integrals can be estimated safely by replacing the finite limits by $-\infty$ ($+\infty$) and additionally by approximating the terms $1-y/\langle l\rangle\approx1$. This yields $\lim_{n\to\infty}\mathcal{T}_0=1$ and $\lim_{n\to\infty}\mathcal{T}_1=0$. Hence,

$$H_n\overset{n\to\infty}{=}-\frac{n}{\langle l\rangle}H[\mathbf{p}]. \tag{39}$$

This is exactly what was obtained by intuitive reasoning [26].

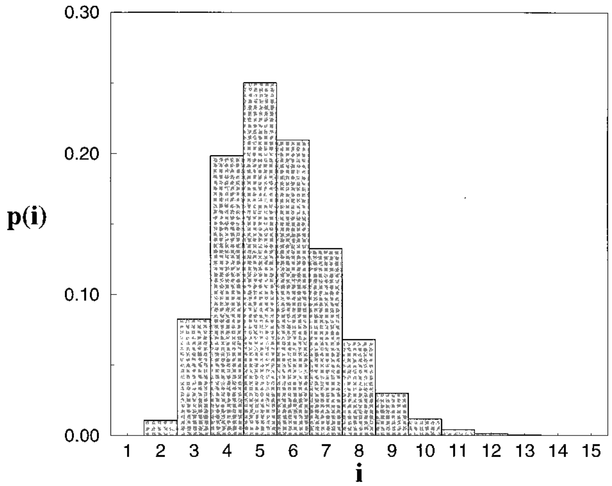

FIG. 1. A histogram plot of the chosen laminar length distribution (41) with resulting values $\langle l\rangle \approx 5.5$, $\sigma_{ll} \approx 2.75$, $H[\mathbf{p}] \approx 2.73$, $A[\mathbf{p}] \approx -0.26$.

An analytic evaluation of the integrals (37) and (38) is not possible and most basic approximations which render the integrals calculable eventually sacrifice the slight deviations from the simple Gaussian form, hence destroying the whole effect. This is even worse when intending to calculate the h_n since then we have to subtract two quantities, H_{n+1} and H_n, which differ only slightly. We regard formula (36) supplemented by (37) and (38) as our final result. In the next section we will exemplify the theoretical result by evaluating the integrals (and all other ingredients) numerically.

III. NUMERICAL EVALUATION AND SIMULATION

The choice of an example is restricted by the fact that we want the asymptotic region, i.e., $n \gg \langle l\rangle$, not to be too far out. On the other hand, it should not be too simple in order to see some deviation from a simple Bernoulli process. We expect a rapid decrease of the h_n approaching the limit

$$h = \frac{H[\mathbf{p}]}{\langle l\rangle} \tag{40}$$

for values around $\langle l\rangle$. The final decay of the h_n for $n > \langle l\rangle$ should be described by our theoretical formula (36).

The statistics of the intermittent sequence we consider is governed completely by the distribution of laminar lengths $p(i)$. We choose

$$p(i) = \frac{i^{10}\exp(-2i)}{\Sigma_{j=1}^{15} p(j)}. \tag{41}$$

A histogram plot of this distribution is given in Fig. 1.

The ingredients for our formula (36) can be calculated and read $\langle l\rangle \approx 5.5$, $\sigma_{ll} \approx 2.75$, $H[\mathbf{p}] \approx 2.73$, $A[\mathbf{p}] \approx -0.26$. This choice is supposed to mimic the length distribution of words in a realistic text.

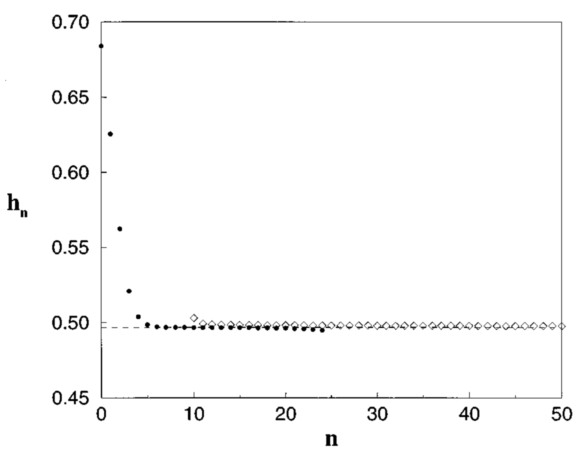

FIG. 2. The conditional entropies for the length distribution (41) in the range $n = 0,1,\ldots,50$; black circles: string simulation (sample length 10^7) results for $n = 0,1,\ldots,24$; diamonds: numerically evaluated formula (36) for $n = 10,\ldots,50$. The dashed line shows the limit $h = 0.496\,734$. Note that the finite sample size effect is crucial for simulation data h_n close to the limit h.

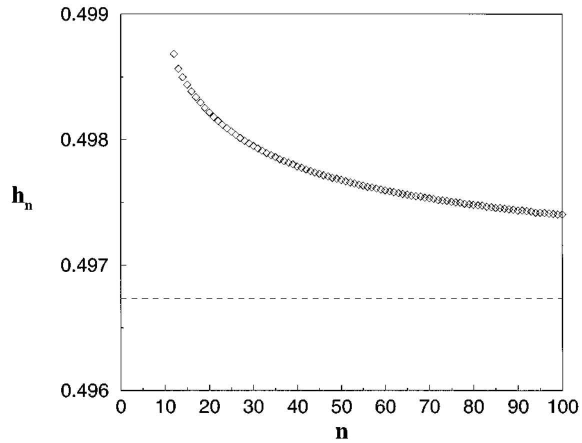

FIG. 3. Enlarged asymptotic decay region as dictated by the numerically evaluated formula (36). A slow decay can be detected. The dashed line shows the limit $h = 0.496\,734$.

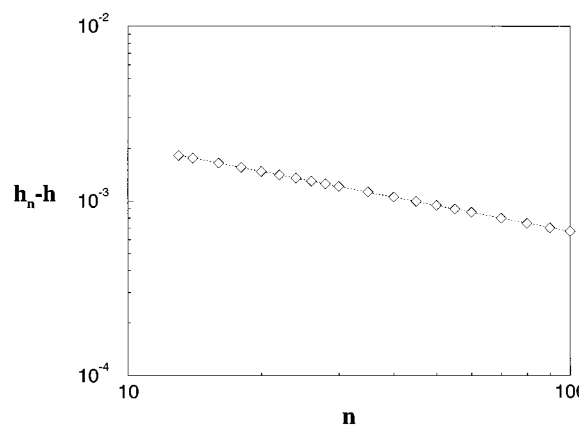

FIG. 4. A double logarithmic representation of $(h_n - h)$ [numerically evaluated formula (36)] yields a straight line. This indicates a power law decay, i.e., $(h_n - h) \sim n^{-\alpha}$. The exponent was obtained by a nonlinear fit yielding $\alpha = 0.492$.

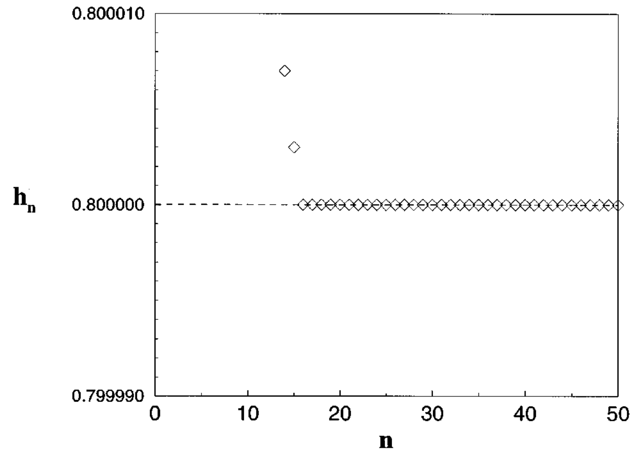

FIG. 5. Decay of conditional entropies for an equidistribution of laminar lengths $p(i)=0.25$ $(i=1,2,3,4)$. The data indicate a ''Markov-like'' behavior (of order 16).

According to (40) we calculate $\lim_{n\to\infty}h_n=h=0.496\,734\ldots$. Figure 2 depicts the h_n for values in the range $n=0,1,\ldots,50$; black circles are the string simulation (sample length 10^7) results for $n=0,1,\ldots,24$; diamonds are the numerically evaluated formula (36) for $n=10,\ldots,50$.

The plot clearly shows the rapid decay for values up to $n\approx\langle l\rangle\approx5$. The numerical values are affected by the finite sample size which generally tends to diminish the values of h_n. Moreover, it can be seen that our analytical approach becomes valid only for $n>\langle l\rangle$. In this asymptotic regime a further decrease of the h_n is hard to detect. For this reason we enlarged the asymptotic decay which can be seen from Fig. 3.

In order to decide the functional character of this decay we subtract the value $h=0.496\,734$ and represent data in a double logarithmic plot; see Fig. 4.

The result is a declining straight line which indicates a power law decay. A nonlinear fit yields

$$(h_n-h)\sim\frac{0.006\,462}{n^{0.492}}. \tag{42}$$

Surprisingly, the asymptotic decay is of a subexponential type.

One could argue that this is a generic effect of the asymptotic formula and not of the system itself. A similar investigation for the distribution $p(i)=\frac{1}{4}$ for $i=1,2,3,4$ yields a different behavior. The representation of the numerically evaluated formula (36) corresponding to Fig. 3 in this case indicates a ''Markov-like'' behavior, i.e., the h_n reach their theoretical limit $h=0.8$ for $n=16$ (see Fig. 5).

Of course, this is no proof since our numerical evaluation is restricted by finite accuracy. At least this question may stimulate further investigations.

IV. CONCLUSION

We have investigated the decay of conditional entropies derived from binary sequences which were related to a regenerative process, i.e., intermittency type I with independent laminar phases. Investigations were performed by string simulations (affected by finite sample size effects) and additionally by applying a theoretical approach valid in the asymptotic regime $n\to\infty$. In accordance with intuition a rapid decay close to the limit value $h=H[\mathbf{p}]/\langle l\rangle$ was observed for values up to $n\approx\langle l\rangle$. In one case the analytic formula hints at a slow decay of h_n in the asymptotic region, whereas in another case a ''Markov-like'' behavior is indicated. A clarifying rigorous result needs further analytical treatment.

ACKNOWLEDGMENT

Discussions with J. Schuchhardt were rather inspiring.

[1] P. Manneville and Y. Pomeau, Phys. Lett. **75A**, 1 (1979).
[2] B. Hu and J. Rudnick, Phys. Rev. Lett. **48**, 1645 (1982).
[3] P. Manneville, J. Phys. (Paris) **41**, 1235 (1980).
[4] H. G. Schuster, *Deterministic Chaos*, 2nd revised ed. (VCH, Weinheim, 1989).
[5] V. S. Anishchenko, *Dynamical Chaos in Physical Systems* (Teubner, Leipzig, 1989).
[6] V. Baladi, J.-P. Eckmann, and D. Ruelle, Nonlinearity **2**, 119 (1989).
[7] X.-J. Wang, Phys. Rev. A **45**, 8407 (1992).
[8] V. M. Alekseev and M. V. Yakobson, Phys. Rep. **75**, 287 (1981).
[9] A. O. Schmitt, H. Herzel, and W. Ebeling, Europhys. Lett. **23**, 303 (1993).
[10] T. Pöschel, W. Ebeling, and H. Rosé, J. Stat. Phys. **80**, 1443 (1995).
[11] A. I. Khinchin, *Mathematical Foundations of Information Theory* (Dover Publ., New York, 1957).
[12] W. Ebeling and G. Nicolis, Chaos Solitons Fractals **2**, 635 (1992).
[13] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).
[14] B. McMillan, Ann. Math. Stat. **24**, 196 (1953).
[15] F.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).
[16] P. Gaspard and X.-J. Wang, Phys. Rep. **235**, 291 (1993).
[17] P. Grassberger, Int. J. Theor. Phys. **25**, 907 (1986).
[18] P. Szépfalusy, Phys. Scr. **T25**, 226 (1989).
[19] L. Gatlin, *Information Theory and the Living System* (Columbia University Press, New York, 1972).
[20] G. Györgyi and P. Szépfalusy, Phys. Rev. A **31**, 3477 (1985).
[21] P. Szépfalusy and G. Györgyi, Phys. Rev. A **33**, 2852 (1986).
[22] A. Csordás, G. Györgyi, P. Szépfalusy, and T. Tél, Chaos **3**, 31 (1993).
[23] W. Hilberg, Frequenz **44**, 243 (1990).
[24] W. Ebeling and G. Nicolis, Europhys. Lett. **14**, 191 (1991).
[25] W. Ebeling and T. Pöschel, Europhys. Lett. **26**, 241 (1994).
[26] J. Freund and H. Herzel, Chaos Solitons Fractals **7**, 49 (1996).
[27] A. Rényi, *Probability Theory* (North-Holland, Amsterdam, 1970).