Expansion-modification systems: A model for spatial $1/f$ spectra

Wentian Li
Santa Fe Institute, 1120 Canyon Road, Santa Fe, New Mexico 87501
(Received 12 September 1990)

It is proposed in this paper that when two processes compete with each other, one tends to create long-range correlations, whereas the other tends to destroy it; the limiting object usually consists of many length scales. In particular, it might have power-law correlation functions and spatial $1/f$ spectra. The basic features of this competing dynamical process with spatial degrees of freedom are captured by the model *expansion-modification* system [W. Li, Europhys. Lett. **10**, 395 (1989)]. Many detailed properties of the system, both numerical and analytical, are discussed here. The implications of the proposed mechanism for power-law correlation functions to the evolution of mass distribution in the Universe and the evolution of prebiotic nucleotide sequences are also discussed.

I. INTRODUCTION

There are many systems that have more than one characteristic length or time scale. For example, the hierarchy of stars, galaxies, cluster of galaxies, and superclusters in the Universe; the nucleotides, codons, genes, and gene clusters in DNA sequences; and the letters, words, sentences, paragraphs, and articles in natural languages. Many names are attached to this type of multiscale (some of them refer to the multiple-time-scale) phenomenon, such as "hierarchy," "fractal,"[1] "$1/f$ noise,"[2] "long tails,"[3] "intermittency,"[4] etc. Some of the names refer to the more restrictive case that the longer-range correlations are scaled with the shorter-range ones, and if some appropriate correlation measures are calculated as the function of distance or as the function of the size of the units, one typically finds power-law functions.

Rather than having a single universal mechanism for all these multiscale phenomena, there are actually many different ways to generate them. The famous example is the critical point in phase transitions.[5] By tuning the temperature to the critical point and observing, say, the two-dimensional pattern of the Ising model, one will find structures with different length scales. Roughly speaking, the multiscale phenomenon at critical point is a result of the competition between the tendencies to be ordered and disordered. The critical point represents the right balance between the two tendencies. If such a balance cannot be reached, the transition from one phase to another does not go through a critical point, and we have the case of the first-order phase transition.

Another well-known mechanism is the breakage process,[6] or the fragmentation, which can be easily illustrated by a process of breaking rocks. The large rocks are broken into small pieces, the small rocks into even smaller ones, and so on. This process is intimately related to the log-normal distribution,[7] a distribution with long tails. Note that there is an upper limit of the largest length scale because the breakage starts from the largest rock. It is important that not all the larger rockes are destroyed, otherwise there are only rocks of smaller sizes. In other words, in order to have rocks with both large and small sizes, there is a tendency to resist the breakage (which can be due to some law of statistics) as well as the breakage itself. An interesting connection between a particular type of fragmentation, the restricted space division, and the $1/f$ spectra is discussed in Refs. 8 and 9. Again, multiscale phenomenon should result from the balance between the two competing factors.

The mechanism to be discussed in this paper reverses the above top-down process to a bottom-up process. Rather than generating smaller scales by breaking down the larger objects, one can build up larger scales by expanding the smaller ones. In order to save the smaller scales, another "destruction" process operates at the same time against the expansion. The combination of the two competing processes can be illustrated by the following idealized model, in which the expansion is written as a discrete equation (symbol 1 is expanded to k 1's at the next time step):

$$1 \rightarrow \underbrace{11\cdots1}_{k} \tag{1.1}$$

and the destruction is modeled by the discrete equation (symbol 1 is "modified" to symbol 0)

$$1 \rightarrow 0 \ . \tag{1.2}$$

Similar rules also apply to symbol 0 to maintain the symmetry between 0 and 1. This set of rules and other variations are called *expansion-modification* systems, first discussed in Ref. 10.

This paper will discuss many properties of the expansion-modification systems. Section II introduces the simplest expansion-modification system with each symbol expanded to two symbols. Section III shows that the correlation functions for sequences generated by expansion-modification systems are typically power-law function, and derives the exponent of this power-law function. Section IV shows that not all rewriting rules that increase the sequence length can generate power-law correlation functions. Section V discusses other

expansion-modification systems, especially those that always rewrite one symbol into two symbols. Section VI presents an attempt to apply the expansion-modification mechanism to cellular automata. Section VII discusses the implication of the model to the evolution of the Universe and the evolution of nucleotide sequences.

Appendixes include some mathematical details used in this paper. Appendix A (to be used in Sec. II) calculates correlation function for sequences randomly packed with blocks. Appendix B (to be used in Secs. III and IV) derives the transition probabilities relating the correlation function at two consecutive time steps. Appendix C (to be used in Sec. V) derives the eigenvalues of the transition matrix for all two-symbol Lindenmayer systems, which always rewrite each symbol to two symbols.

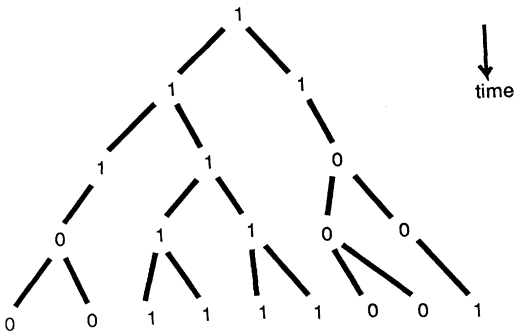

FIG. 1. A particular realization of the rewriting process for the simplest expansion-modification system (2.1).

II. THE SIMPLEST EXPANSION-MODIFICATION SYSTEM

The simplest expansion-modification system is the two-symbol system in which the expansion process rewrites one symbol to two identical symbols and the modification process switches one symbol to another symbol, i.e.,[10]

$$\begin{aligned} 1 &\to \begin{cases} 11 & \{1-q\} \\ 0 & \{q\} \end{cases}, \\ 0 &\to \begin{cases} 00 & \{1-p\} \\ 1 & \{p\} \end{cases}. \end{aligned} \tag{2.1}$$

The parameter values in curly brackets are the probabilities for the rewriting rules to be applied. p and q are modification probabilities, or, to borrow a terminology from biology, *mutation rates.* I almost always set $p=q$ in this paper, so that there is no bias between the two symbols. Most of the conclusions on the behavior of the system will not be affected by this constraint.

By repeated applications of the rewriting rules, a single symbol expands to a sequence, and a shorter sequence becomes a longer sequence. Denoting $P_1(t)$ for the density of symbol 1 at time t, the average length of the sequence at time t is $\overline{N(t)}=k(t)k(t-1)k(t-2)\cdots k(0)\overline{N(0)}$, where

$$\begin{aligned} k(t) &= (2-q)P_1(t)+(2-p)[1-P_1(t)] \\ &= 2-q+(q-p)P_1(t) \end{aligned} \tag{2.2}$$

is the average *branching ratio* at time t. When $p=q$, k is equal to $2-p$, which is the same constant at all times. In general, if $k=\lim_{t\to\infty}k(t)$, called the *limiting branching ratio,* the sequence length $\overline{N(t)}\sim k^t\overline{N(0)}$ increases exponentially. The rewriting process represented by Eq. (2.1) is illustrated in Fig. 1.

This type of rewriting rule belongs to what is called the *probabilisitic context-free Lindenmayer system*.[11] Lindenmayer systems are grammars (in the framework of formal language theory[12]), which update symbols synchronously. The name context-free means that there is no coupling among neighboring symbols in a sequence during the rewriting process. Note that the context-free language studied in formal language theory[12] is different from a context-free Lindenmayer system in two aspects: (i) it has another set of symbols (nonterminal symbols) which do not appear in the final sequence; and (ii) it does not necessarily need synchronous updating. The second difference means that for context-free languages, different branches of the tree (similar to what is shown in Fig. 1) can stop growing at different times, whereas for context-free Lindenmayer systems, they all stop simultaneously.

Lindenmayer systems can also be considered as dynamical systems with spatial degrees of freedom. From this perspective, they can be called "open" dynamical systems[10] in comparison with those whose number of degrees of freedom is fixed during the dynamics. Actually, the term "open" has an even wider meaning[13] that the rule itself can be subject to variations by the environment or by some higher-level instructions (but they are usually fixed). Lindenmayer systems are not "open" by the second definition because the rewriting rule is fixed.

In order to see how the competition between expansion and modification affect the statistical properties of the limiting sequences, in the following, I examine three extreme cases: (i) $p=0$, no modification; (ii) $p=1$, no expansion; and (iii) $p=0.5$, the expansion and the modification is applied with equal probability.

When $p=0$, the limiting sequence is exactly the expanded version of the initial sequence. For example, if the initial sequence is a single symbol (e.g., 1), the limiting sequence is a string of 1's; if the initial sequence is 10, the limiting sequence contains half 1's and half 0's. An interesting case is when the initial sequence is a random sequence; the limiting sequence is then a random packing of blocks. After the rewriting rule is applied t times, the length for each block is 2^t, and the joint probability for having two 1's separated by distance d, is (see Appendix A)

$$P_{11}(d)=\frac{N_{11}[i-1]}{N}(i2^t-d)+\frac{N_{11}[i]}{N}[d-(i-1)2^t] \quad \text{if } (i-1)2^t\le d<i2^t \tag{2.3}$$

where $N_{11}[i]$ is the number of 1-to-1 pairs separated by distance i at time 0, and N is the sequence length. The correlation function $\Gamma(d)=P_{11}(d)-P_1^2$ behaves the same as $P_{11}(d)$, except for a constant term, so $\Gamma(d)$ is a piece-

wise linear function.

When $p=1$, the dynamics is periodic and the limiting sequence oscillates between the copy and the complementary copy of the initial sequence. No new structures or new correlations are generated by the rewriting.

When $p=0.5$ or $p \geq 0.5$, the mutation rate is large enough to destroy most of the long-range correlation in the limiting sequence. Although there are always short-range correlations being generated that lead to nonwhite spectra at higher frequencies, the limiting sequence is nevertheless very close to a random sequence.

The most interesting range for the mutation rate is when it is small but nonzero, such as $0<p<0.2$. In this case, the expansion part of the rule has a long period of time to generate larger blocks without disruption, but eventually a mutation will cut a block into two, and the cutting point produced by mutation keeps growing itself, and so on. Intuitively, the self-similar dynamics will be reflected by the self-similarity in the limiting sequence, for example, the power-law decay of correlation functions and power spectra. (Although a self-similar dynamics itself cannot guarantee the limiting sequence to have power-law correlations as in the case of p being large). Figure 2 shows the spatial-temporal patterns for rule (2.1) at (a) $p=0.01$, (b) 0.1, (c) 0.5, and (d) 0.9.

In order to verify this, the numerical calculation of the power spectrum is carried out for the limiting sequences at different mutation rates. The power spectrum $P(f)$ is the Fourier transformation of the correlation function $\Gamma(d)$, or the square of the Fourier amplitudes $A(f)$ of the sequence $\{x_j\}$:

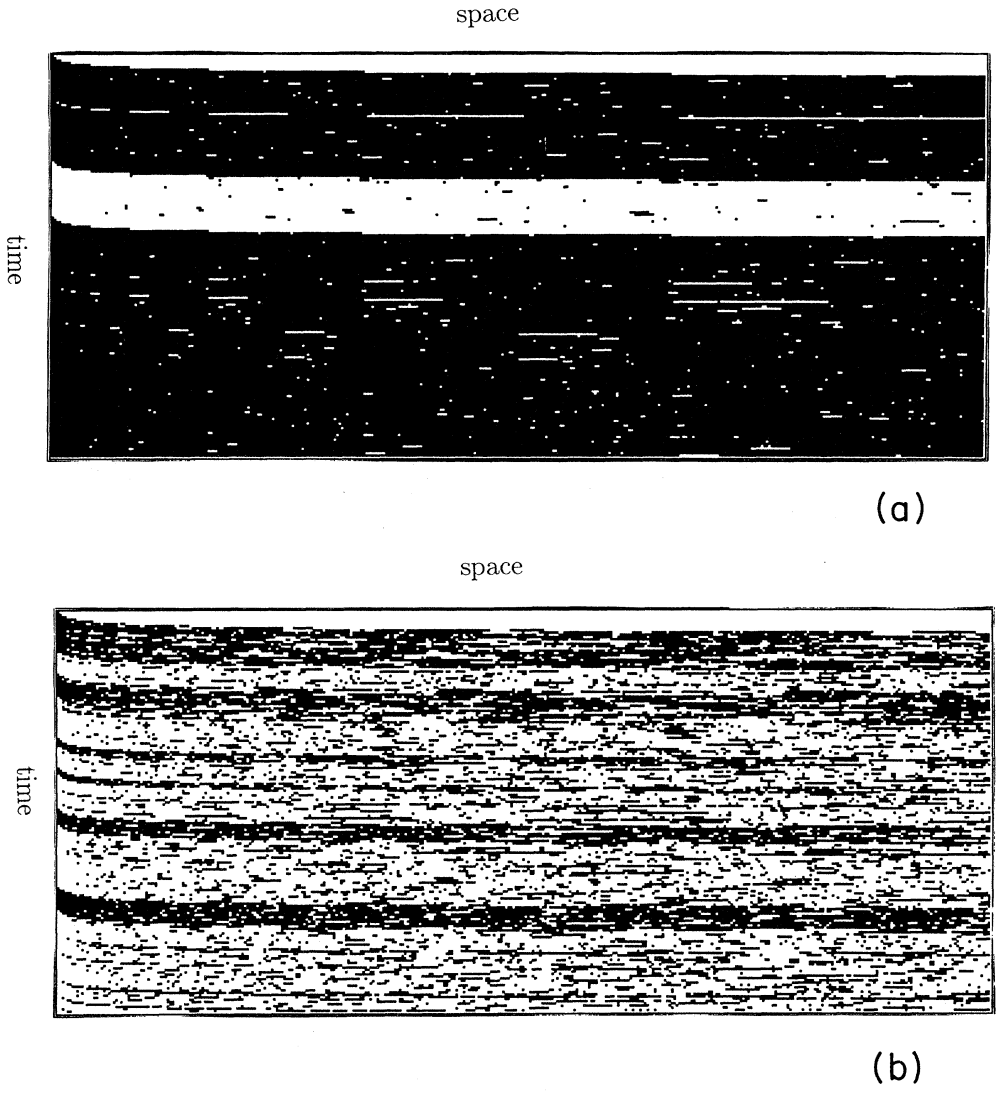

FIG. 2. Spatial-temporal patterns of the simplest expansion-modification system (2.1) at the mutation rate (a) $p=0.01$, (b) 0.1, (c) 0.5, and (d) 0.9. q is equal to p. The sequences at different time steps are aligned to the left by their first sites. The number of time steps is 175. The cutoff of the sequence length is 380. If the sequence length is shorter than 380, the remaining space is left blank. If the sequence length is longer than 380, only the first 380 sites are shown. The time arrow is pointing down.

$$A(f)=\frac{1}{N}\sum_{j=0}^{N-1} x_j e^{-i2\pi(f/N)j} \quad (f=0,\ldots,N-1)\ , \qquad (2.4)$$

i.e., $P(f)=N|A(f)|^2$. By simple scaling property of the Fourier transformation (see, e.g., Chap. 12 of Ref. 14), it can be shown that a power-law power spectrum $P(f)\sim 1/f^{\alpha}$ is equivalent to a power-law correlation function $\Gamma(d)\sim 1/f^{1-\alpha}$. For example, the "$1/f$ noise" corresponds to the extremely slowly decaying correlation function (with the exponent $1-\alpha$ almost equal to zero), and the "$1/f^2$ noise" corresponds to linear decay of the correlation function. The example given above (at $p=0$) of the sequence with randomly packed blocks has a $1/f^2$ spectrum because its correlation function [Eq. (2.3)] is linear.

Figure 3 shows the power spectra (in log-log scales) of the limiting sequences generated at (a) $p=0.01$, (b) 0.1, (c) 0.5, and (d) 0.9. The limiting sequences are taken at the time whenever the sequence length becomes longer than $2^{13}=8192$. At $p=0.01$ and 0.1, the power-law function $1/f^{\alpha}$ can fit the spectra perfectly, with α very close to 1. It is amazing of how easy a $1/f$ spectrum can be generated in this way.

This $1/f$ spectra cannot persist for all parameter values of p, as we know that by gradually increasing the mutation rate beyond, say, 0.5, the limiting sequence is almost random and the power spectrum should be almost flat, which implies α is zero instead of 1. Therefore there should be a transition from the $1/f$ spectrum to the white spectrum.

More careful numerical studies show that the exponent of the power-law function for the main scaling region (lower-frequency part) of the power spectrum does indeed decrease as p is increased. Actually, the spectrum is not always a good power-law function. The scaling remains at lower frequencies, but bend over at higher frequencies. From the lower frequency scaling region to higher frequency tail, there is a small crossover region which more or less stays as $1/f$, although it can be hard to judge whether or not the crossover is a continuous one.

Figure 4 shows how the exponents ($-\alpha$ in $f^{-\alpha}$) change with p. Two sets of data points are included in order to show that the lower-frequency part and the middle-frequency part have different exponents. Actually, the spectra can roughly be partitioned into three parts. I did

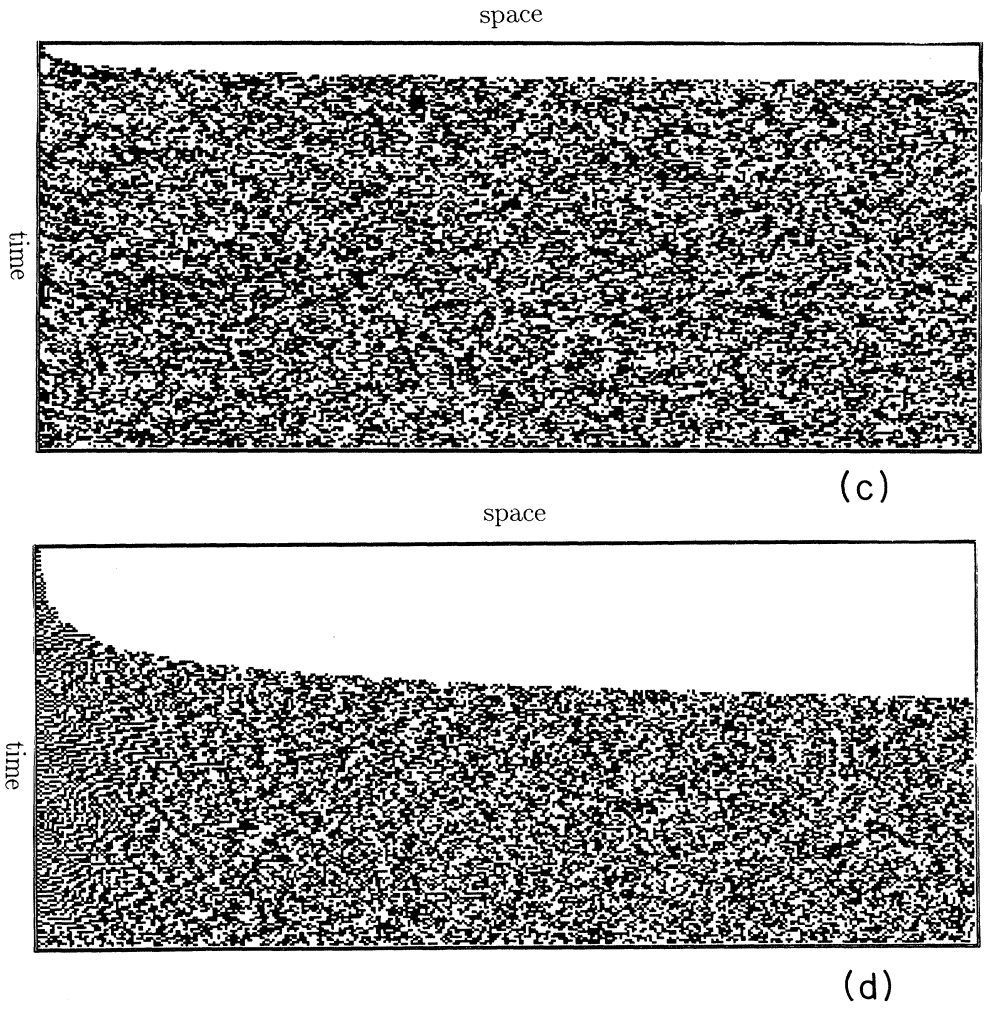

FIG. 2. (*Continued*).

not include the exponent for the high-frequency tail because it is only related to the short-range correlation. Section III will show the analytic results, which will explain why the correlation functions and, consequently, power spectra are power-law functions, and how the exponent (the one for the low-frequency part) changes with p.

III. EXPONENT α OF THE $1/f^{\alpha}$ SPECTRA

In this section, I will give an explanation as to why the correlation functions of the limiting sequences for rule Eq. (2.1) decay according to the power law, and derive the exponent for this power-law function. Generally speaking, the limiting sequences of context-free Lindenmayer systems have power-law correlation functions if the largest nontrivial eigenvalue of the transition matrix (to be explained later) is positive. The case for context-free Lindenmayer systems with negative eigenvalues of their transition matrix will be discussed in Sec. IV.

Suppose that we look at the sequence after the rule is applied t times: $\{x_i^t\}$ $[i=1,2,\ldots,N(t)]$; each site value can be one of the n symbols $\{a_\alpha\}$ $(\alpha=1,2,\ldots,n)$. As a convention, i, j, and k are used for the site indices and α, β, and γ for symbol state indices. The joint probability for having the symbol pair α-β separated by the distance d is $[\mathcal{P}_{\alpha\beta}(d)]^t$, with the superscript indicating the time. Upon applying the rewriting rule, this symbol pair α-β will lead to other symbol pairs α'-β' separated by longer distances d' at the next time step. Assuming that the transition probability from an α-β pair to an α'-β' pair is $T(\alpha\beta d\rightarrow\alpha'\beta' d')$, the joint probability satisfies the following dynamical equation:

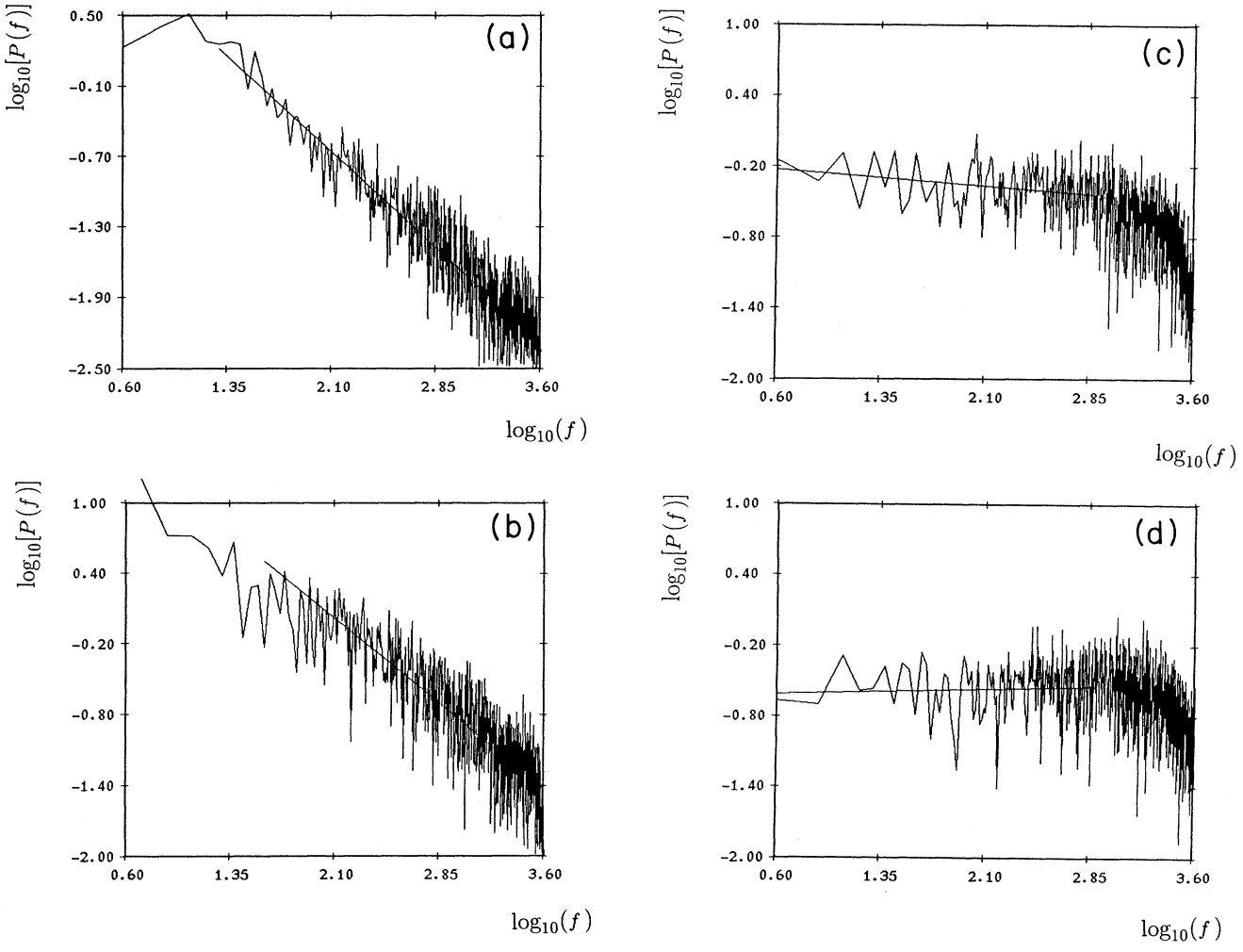

FIG. 3. Spatial power spectra $P(f)$ (in $\log_{10}$-$\log_{10}$ scales) for sequences generated by the simplest expansion-modification system (2.1). The rewriting is started from a single symbol 1, and stopped whenever the sequence length is longer than $2^{13}=8192$. Half of the fast Fourier transformation (FFT) components are redundant, and the neighboring four components are averaged as one component. It leaves $2^{10}=1024$ points in the plot. (a) $p=0.01$, a fitting straight line has slope -1.1, or it is the function $1/f^{1.1}$; (b) $p=0.1$, a fitting straight line has slope -0.96; (c) $p=0.5$, with a straight line having slope -0.09; and (d) $p=0.9$, with a straight line having a slope 0.03.

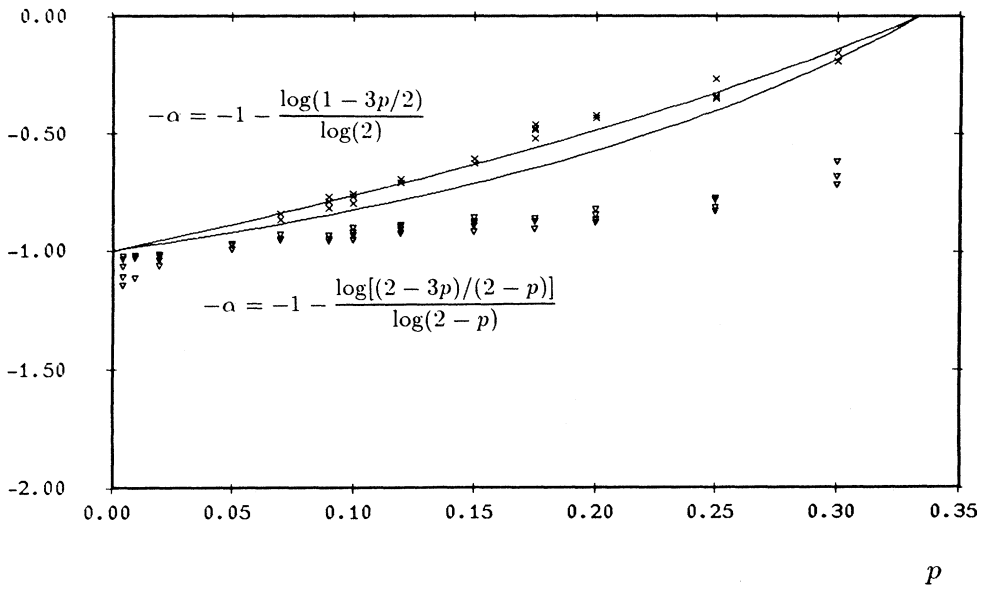

FIG. 4. Numerically estimated exponents $-\alpha$ for the power-law function approximation $1/f^{\alpha}$ of the spectra in Fig. 3, as a function of mutation rate p. The values represented by crosses are for lower-frequency parts of the spectra; those represented by triangles are for the middle parts. The top line is the analytic estimation of the exponent α by Eq. (5.11) for the fixed-branching-ratio-2 rule (5.2). Another line is the approximated estimation of the exponent α by Eq. (3.18) for the simplest expansion-modification rule (2.1).

$$[\mathcal{P}_{\alpha'\beta'}(d')]^{t+1}=\sum_{d}\sum_{\alpha,\beta}T(\alpha\beta d\rightarrow\alpha'\beta'd')[\mathcal{P}_{\alpha\beta}(d)]^{t}\ . \tag{3.1}$$

If the $T(\alpha\beta d\rightarrow\alpha'\beta'd')$ are written in matrix form, it is called the *transition matrix* for the context-free Lindenmayer system.

Suppose there is a time invariant condition in the $t\rightarrow\infty$ limit (if the rewriting leads to periodic or chaotic dynamics of the joint probabilities, this argument does not apply), and the superscript is dropped. The joint probabilities are then the solution of the invariant equation

$$\mathcal{P}_{\alpha'\beta'}(d')=\sum_{d}\sum_{\alpha,\beta}T(\alpha\beta d\rightarrow\alpha'\beta'd')\mathcal{P}_{\alpha\beta}(d) \tag{3.2}$$

or in a form of the matrix equation

$$\mathcal{P}(d')=\sum_{d}\underline{T}(d,d')\mathcal{P}(d)\ , \tag{3.3}$$

where the vector $\mathcal{P}(d)$ contains all the joint probabilities $\mathcal{P}_{\alpha\beta}(d)$ and the matrix $\underline{T}(d,d')$ contains all the transition probabilities $T(\alpha\beta d\rightarrow\alpha'\beta'd')$.

To see what the typical functional solution is, we assume that (i) the matrix equation is approximated by a scalar equation, i.e., the summation over α,β is dropped, and the transition matrix is replaced by the largest nontrivial eigenvalue $\lambda(d,d')$; and (ii) the typical expansion of the distance is the same with the expansion of the whole sequence: $d'/d\approx N(t+1)/N(t)=k$, the the multiscaling relation is approximated by a single scaling relation, i.e., the summation over d is dropped. We then have the approximate equation

$$P(d')=\lambda(d'/k,d')P(d'/k)\ . \tag{3.4}$$

Suppose $\lambda(d'/k,d')=\lambda>0$, the power-law function $P(d)=1/d^{c}$ is a solution of this equation, and the exponent c is

$$c=-\frac{\log(\lambda)}{\log(k)}\ . \tag{3.5}$$

The above argument is general and always leads to a power-law solution as long as λ is a positive constant. In other words, the sequences generated from top to bottom without interactions among neighboring sites will have power-law decay correlation functions if the largest nontrivial eigenvalue of the transition matrix is positive and does not change very much with the distance d. It is in contrast with the case where sequences are generated from left to right with only short memories, such as the Markov chains[15] and the regular languages.[12] In those cases, the correlation functions typically decay exponentially.[15,16] For sequences generated from top to bottom with neighboring couplings such as cellular automata[17] and context-sensitive Lindenmayer systems, there is no generic form for the correlation function because the statistical properties of the limiting sequence are crucially determined by the rule of the interaction.[16]

If we relax the single scaling relation but assume that most of the contribution in the summation over d in Eq. (3.1) is from a few distances around the most likely distance $d=d'/k$, after inserting a power-law solution into the equation, we have

$$\frac{1}{d'^{c}}=\sum_{i(\geq 0)}\lambda\left[\frac{d'}{k}\pm i,d'\right]\frac{1}{(d'/k\pm i)^{c}}$$
$$\approx\sum_{i(\geq 0)}\lambda\left[\frac{d'}{k}\pm i,d'\right]\frac{k^{c}}{d'^{c}}\ , \tag{3.6}$$

then

$$c \approx -\frac{\log\left[\sum_i \lambda(d'/k \pm i, d')\right]}{\log(k)} . \tag{3.7}$$

Again, if the exponent c calculated by Eq. (3.7) is a function of d', the solution would not be a power-law function with a fixed exponent. The above formula for the exponent c is only an approximation. In general, the multiparameter scaling might be poorly approximated by a scaling with a single parameter. Similar reason leads people to generalize the fractal description to the multifractal description.[18]

To illustrate the above discussion, I now calculate the exponent c for the simplest expansion-modification system discussed in Sec. II. First of all, the transition probability $T(\alpha\beta d \to \alpha'\beta' d')$ can be grouped into three types: (i) T_0, which keep both symbol unchanged ($\alpha=\alpha', \beta=\beta'$); (ii) T_1 which change one symbol (either $\alpha\neq\alpha'$ or $\beta\neq\beta'$); and (iii) T_2, which change both symbols ($\alpha\neq\alpha'$ and $\beta\neq\beta'$). They are (see Appendix B for the derivation)

$$\begin{aligned} T_0 &\equiv T(\alpha\beta d \to \alpha\beta d') \\ &= \frac{(1-p)^2}{2-p}[B_{m+2}(d-1,p) \\ &\qquad + 2B_{m+1}(d-1,p) + B_m(d-1,p)] , \\ T_1 &\equiv T(\alpha\beta d \to \bar{\alpha}\beta d') \\ &= \frac{p(1-p)}{2-p}[B_{m+1}(d-1,p) + B_m(d-1,p)] , \\ T_2 &\equiv T(\alpha\beta d \to \bar{\alpha}\bar{\beta} d') = \frac{p^2}{2-p} B_m(d-1,p) , \end{aligned} \tag{3.8}$$

where the overhead bar represents the switching operation between the two symbols. m is roughly the number of sites in between the α-β pair which mutate

$$m \equiv (d-1)-(d'-d) = 2d-d'-1 , \tag{3.9}$$

$d'-d$ is the number of sites which expand, and $B_m(n,p)$ is the binomial coefficient

$$B_m(n,p) \equiv \begin{bmatrix} n \\ m \end{bmatrix} p^m (1-p)^{n-m} , \tag{3.10}$$

which is the probability for one event, with probability p, to occur m times and another event, with probability $1-p$, to occur $n-m$ times. Because of the symmetry, $T(\alpha\beta d \to \alpha\bar{\beta} d')$ is equal to $T(\alpha\beta d \to \bar{\alpha}\beta d') = T_1$.

The way to remember these formulas is to note that whenever the expansion is applied to both symbol α and β, there is a factor of $(1-p)^2$; if there are two mutations, the coefficient contains p^2. As for the subscripts in the binomial coefficients, there is $m+2$ whenever the transition is through the two outer legs of the expansions; $m+1$ if the transition is through one outer leg and one inner leg of two expansions, or one outer leg of expansion and one mutation; and m if the transition is through one inner leg of the expansion and one mutation, or two mutations (see Figs. 9 and 10 for an illustration).

Due to the normalization condition of the probabilities, we have

$$\begin{aligned} &\sum_{d_{\min} \le d \le d_{\max}} [T_0 + 2T_1 + T_2] \\ &\quad = \sum_d \left[\frac{(1-p)^2}{2-p} B_{m+2} + \frac{2(1-p)}{2-p} B_{m+1} + \frac{1}{2-p} B_m \right] \\ &\quad = 1 , \end{aligned} \tag{3.11}$$

where $d_{\min}$ is either $d'/2$ (if d' is even) or $(d'-1)/2$ (if d' is odd), and $d_{\max} = d'$. These are situations when all sites (or almost all sites) are doubled ($d = d_{\min}$) and all are mutated ($d = d_{\max}$), respectively.

The transition matrix is of the form

$$\underline{T}(d,d') = \begin{bmatrix} T_0 & T_1 & T_1 & T_2 \\ T_1 & T_0 & T_2 & T_1 \\ T_1 & T_2 & T_0 & T_1 \\ T_2 & T_1 & T_1 & T_0 \end{bmatrix} \tag{3.12}$$

(all the matrix elements are function of d, d', and the mutation rate p) and its eigenvalues are

$$\lambda_1(d,d') = T_0 + 2T_1 + T_2 , \tag{3.13}$$

$$\lambda_2(d,d') = \lambda_3(d,d') = T_0 - T_2 , \tag{3.14}$$

$$\lambda_4(d,d') = T_0 - 2T_1 + T_2 . \tag{3.15}$$

$\lambda_1(d,d')$ is the "trivial" eigenvalue because $\sum_d \lambda_1(d,d') = 1$. In the parameter range of interest ($0 < p < 0.5$), $\lambda_2(d,d')$ is larger than $\lambda_4(d,d')$ (all eigenvalues are positive). So $\lambda_2(d,d')$ is the largest nontrivial eigenvalue which dominates the scaling behavior of the correlation function. We can calculate the summation of $\lambda_2(d,d')$ over d:

$$\begin{aligned} \sum_d \lambda_2(d,d') &= \sum_d (T_0 - T_2) \\ &= \sum_d (T_0 + 2T_1 + T_2) - \sum_d 2(T_1 + T_2) \\ &= 1 - \frac{2p}{2-p} = \frac{2-3p}{2-p} , \end{aligned} \tag{3.16}$$

and using Eq. (3.7), the exponent c as a function of p is

$$c \approx -\frac{\log[(2-3p)/(2-p)]}{\log(2-p)} = 1 - \frac{\log(2-3p)}{\log(2-p)} . \tag{3.17}$$

Finally, it is straightforward to determine the exponent α in the power spectrum $P(f) \sim 1/f^{\alpha}$:

$$\alpha = 1 - c \approx 1 + \frac{\log[(2-3p)/(2-p)]}{\log(2-p)} = \frac{\log(2-3p)}{\log(2-p)} . \tag{3.18}$$

In the limit of $p \to 0$, we have $c \to 0$ and $\alpha \to 1$. In other words, we have a $1/f$ spectrum. The result is plotted in Fig. 4 in comparison with the numerically determined exponent. They fit quite well.

IV. CONTEXT-FREE LINDENMAYER SYSTEMS THAT CANNOT GENERATE $1/f^{\alpha}$ SPECTRA

In Sec. III, it has been shown that context-free Lindemayer systems typically have limiting sequences with power-law decay correlation functions. It is the case when the largest nontrivial eigenvalue of the transition matrix is positive and the exponent c calculated by Eq. (3.7) does not depend on the distance d. As will be seen in this section, these conditions can be violated for other context-free Lindenmayer systems.

Considering the following rule, which is similar to the expansion-modification system (2.1) except that the expansion is replaced by a rewriting to two different symbols instead of two same symbols,

$$\begin{aligned} 1 &\to \begin{cases} 10 & \{1-q\} \\ 0 & \{q\} \end{cases}, \\ 0 &\to \begin{cases} 01 & \{1-p\} \\ 1 & \{p\} \end{cases}. \end{aligned} \tag{4.1}$$

If not specially claimed, q is set to be equal to p.

In the $p=1$ limit, rule (4.1) is the same as rule (2.1), and the limiting sequence oscillates between the direct and the complementary copies. At $p=0$, the limiting sequence is the Thue-Morse sequence,[19,20] which is almost periodic but not exactly. The correlation function for almost-periodic sequences are oscillating functions, and should not be a monotonically decaying power-law function as in the case in Sec. III. This suggests that the largest eigenvalue (in absolute value) of the transition matrix for Eq. (4.1) should be negative.

To check this, I write down the transition probabilities $T(\alpha\beta d \to \alpha'\beta' d')$ for Eq. (4.1) (see Appendix B for the derivation):

$$\begin{aligned} T_0 &\equiv T(\alpha\beta d \to \alpha\beta d') = \frac{(1-p)^2}{2-p} B_{m+1}(d-1,p), \\ T_{11} &\equiv T(\alpha\beta d \to \bar{\alpha}\beta d') = \frac{1-p}{2-p} B_m(d-1,p), \\ T_{12} &\equiv T(\alpha\beta d \to \alpha\bar{\beta} d') \\ &= \frac{(1-p)^2}{2-p} B_{m+2}(d-1,p) \\ &\quad + \frac{(1-p)p}{2-p} B_{m+1}(d-1,p), \\ T_2 &\equiv T(\alpha\beta d \to \bar{\alpha}\bar{\beta} d') \\ &= \frac{1-p}{2-p} B_{m+1}(d-1,p) + \frac{p}{2-p} B_m(d-1,p), \end{aligned} \tag{4.2}$$

Unlike the Lindenmayer system (2.1), here the probability of changing one symbol depends on whether the left or the right symbol is changed, so $T_{11} \neq T_{12}$.

The transition matrix can be written as

$$\underline{T}(d,d') = \begin{bmatrix} T_0 & T_{12} & T_{11} & T_2 \\ T_{12} & T_0 & T_2 & T_{11} \\ T_{11} & T_2 & T_0 & T_{12} \\ T_2 & T_{11} & T_{12} & T_0 \end{bmatrix}, \tag{4.3}$$

and its eigenvalues are

$$\lambda_1(d,d') = T_0 + T_{11} + T_{12} + T_2, \tag{4.4}$$

$$\lambda_2(d,d') = T_0 + T_{11} - T_{12} - T_2, \tag{4.5}$$

$$\lambda_3(d,d') = T_0 - T_{11} + T_{12} - T_2, \tag{4.6}$$

$$\lambda_4(d,d') = T_0 - T_{11} - T_{12} + T_2. \tag{4.7}$$

As in Sec. III, it can be shown that $\sum_d \lambda_1(d,d') = 1$, so $\lambda_1(d,d')$ is the trivial largest eigenvalue. Both λ_2 and λ_3 are negative, whereas $\lambda_4 > 0$. It can be shown that the absolute values of λ_2 and λ_3 are larger than λ_4, and λ_2 is almost equal to λ_3. One can take either λ_2 or λ_3 as the largest (in absolute value) nontrivial eigenvalue.

What is the functional solution for the scaling equation

$$P_{\alpha\beta}(d') = -|\lambda| P_{\alpha\beta}(d'/k) \tag{4.8}$$

with a negative scaling coefficient? If the modulated power-law function is used,

$$P_{\alpha\beta}(d') = e^{i2\pi\theta(d')} \frac{1}{d'^c}, \tag{4.9}$$

we have $c = -\log|\lambda|/\log(k)$ as before, but there is another relation for the phase

$$\theta(d') = \theta(d'/k) + 1/2 + n, \tag{4.10}$$

where n is any integer. One solution of the equation is

$$\theta(d') = \frac{\log d'}{2\log k}. \tag{4.11}$$

In other words, the modulation term changes with distances. An oscillating correlation function leads to peaks at nonzero frequencies in the power spectrum (see, e.g., Ref. 16). It is obvious that it is no longer a $1/f^{\alpha}$ spectrum.

V. OTHER EXPANSION-MODIFICATION SYSTEMS

The simplest expansion-modification system discussed in Sec. II captures most of the spirit of the mechanism involving two competing forces. Naturally, one can make many variations; for example, rather than a symbol being expanded to two identical symbols, it can expand to three identical symbols:

$$\begin{aligned} 1 &\to \begin{cases} 111 & \{1-q\} \\ 0 & \{q\} \end{cases}, \\ 0 &\to \begin{cases} 000 & \{1-p\} \\ 1 & \{p\} \end{cases}. \end{aligned} \tag{5.1}$$

Similar to that in Sec. III (and Appendix B), the procedure can be repeated to derive the transition probabilities and the exponent of the power-law correlation function. This derivation is more complicated though, because there are more configurations to be counted. Here I will only present a numerical result of the power spectrum for the limiting sequence of Eq. (5.1) with $p=0.1$, as shown in Fig. 5, to illustrate that minor changes in the rule would not destroy the basic feature, i.e., the concept

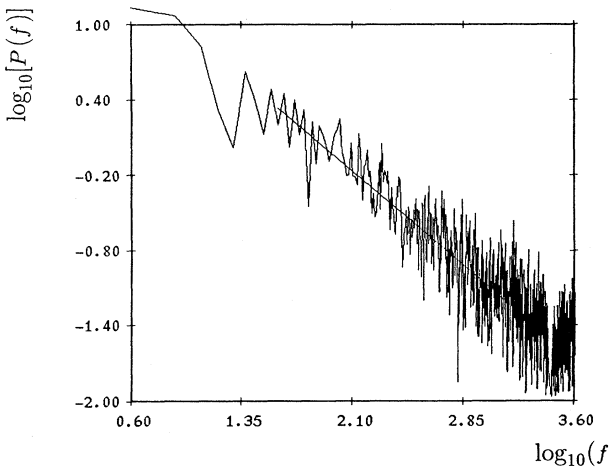

FIG. 5. The spatial power spectrum $P(f)$ (in $\log_{10}$-$\log_{10}$ scale) for the sequence generated by "expansion by three" rule (5.1), at mutation rate $p=0.1$. The sequence length is $2^{13}=8192$. Also shown is a fitting straight line with the slope -1.

of "universality class." The spectrum has the same $1/f$ scaling, and it differs from the spectrum in Fig. 3(b) only in that there is a dip at $f/N=\frac{1}{3}$ (N is the sequence length).

Another variation of the Eq. (2.1) is to add extra symbol (e.g., symbol 0) for the mutation part of the rule:

$$\begin{aligned} 1 &\to \begin{cases} 11 & \{1-q\} \\ 00 & \{q\} \end{cases}, \\ 0 &\to \begin{cases} 00 & \{1-p\} \\ 10 & \{p\} \end{cases}. \end{aligned} \tag{5.2}$$

This type of Lindenmayer system can be called the *fixed-branching-ratio* Lindenmayer system because each instruction of the rule always rewrites one symbol to a fixed number (e.g., 2) of other symbols.

The transition probabilities for fixed-branching-ratio Lindenmayer systems are much easier to calculate because the relation between d' and d is fixed, and no binomial coefficient is involved. For the case of a fixed branching ratio of 2, there are only three possible relations between d' and d: (i) $d'=2d$, (ii) $d'=2d-1$, and (iii) $d'=2d+1$.

If I use $T_L(\alpha\to\alpha')$ to represent the probability from symbol α to α' through the left leg and $T_R(\alpha\to\alpha')$ for that through the right leg (see Fig. 9), the transition probabilities for the fixed-branching-ratio-2 Lindenmayer system are

$$\begin{aligned} T\left[\alpha\beta\frac{d'}{2}\to\alpha'\beta'd'\right] &= \frac{T_L(\alpha\to\alpha')T_L(\beta\to\beta')+T_R(\alpha\to\alpha')T_R(\beta\to\beta')}{2}, \\ T\left[\alpha\beta\frac{(d'-1)}{2}\to\alpha'\beta'd'\right] &= \frac{T_L(\alpha\to\alpha')T_R(\beta\to\beta')}{2}, \\ T\left[\alpha\beta\frac{(d'+1)}{2}\to\alpha'\beta'd'\right] &= \frac{T_R(\alpha\to\alpha')T_L(\beta\to\beta')}{2}. \end{aligned} \tag{5.3}$$

For Eq. (5.2), we have

$$\begin{aligned} &T_L(0\to0)=1-p, \quad T_R(0\to0)=1, \\ &T_L(0\to1)=p, \quad T_R(0\to1)=0, \\ &T_L(1\to0)=q, \quad T_R(1\to0)=q, \\ &T_L(1\to1)=1-q, \quad T_R(1\to1)=1-q. \end{aligned} \tag{5.4}$$

Notice the equivalence relation $T(\alpha\beta d'/2\to\alpha'\beta'd')=T(\beta\alpha d'/2\to\beta'\alpha'd')$; the 16 transition probabilities are reduced to 10 independent ones. For rule (5.2), they are

$$\begin{aligned} T(00\to00)&=\frac{(1-p)^2+1}{2}, \\ T(00\to01)&=\frac{(1-p)p}{2}, \\ T(00\to11)&=\frac{p^2}{2}, \\ T(01\to00)&=\frac{(1-p)q+q}{2}, \\ T(01\to01)&=\frac{(1-p)(1-q)+(1-q)}{2}, \\ T(01\to10)&=\frac{pq}{2}, \\ T(01\to11)&=\frac{p(1-q)}{2}, \\ T(11\to00)&=\frac{q^2}{2}, \\ T(11\to01)&=\frac{q(1-q)}{2}, \\ T(11\to11)&=\frac{(1-q)^2}{2}. \end{aligned} \tag{5.5}$$

Write the transition matrix as $T_{2\alpha'+\beta',2\alpha+\beta}=T(\alpha\beta d'/2\to\alpha'\beta'd')$; the eigenvalues of the matrix are

$$\begin{aligned} \lambda_1&=1, \\ \lambda_2=\lambda_3&=1-\frac{p}{2}-q, \\ \lambda_4&=1-p-2q+\frac{p^2}{2}+q^2+pq. \end{aligned} \tag{5.6}$$

It is easy to check that the largest nontrivial eigenvalue is λ_2.

Similarly, we can write down the transition matrix for $T(\alpha\beta(d'+1)/2\to\alpha'\beta'd')$ and $T(\alpha\beta(d'-1)/2\to\alpha'\beta'd')$. There is no symmetry condition among the transition probabilities, so the number of independent ones remains at 16. These two transition matrices are

$$\underline{T}\left[\frac{d'-1}{2},d'\right]=\begin{bmatrix}(1-p)/2 & q(1-p)/2 & q/2 & q^2/2\\ 0 & (1-p)(1-q)/2 & 0 & q(1-q)/2\\ p/2 & pq/2 & (1-q)/2 & (1-q)q/2\\ 0 & p(1-q)/2 & 0 & (1-q)^2/2\end{bmatrix} \tag{5.7}$$

and

$$\underline{T}\left[\frac{d'+1}{2},d'\right]=\begin{bmatrix}(1-p)/2 & q/2 & q(1-p)/2 & q^2/2\\ p/2 & (1-q)/2 & pq/2 & (1-q)q/2\\ 0 & 0 & (1-p)(1-q)/2 & q(1-q)/2\\ 0 & 0 & p(1-q)/2 & (1-q)^2/2\end{bmatrix} . \tag{5.8}$$

The eigenvalues of these two matrices are the same:

$$\begin{aligned}&\lambda_1=1/2\ ,\\ &\lambda_2=(1-q)/2\ ,\\ &\lambda_3=(1-p-q)/2\ ,\\ &\lambda_4=(1-p-2q+pq+q^2)/2\ .\end{aligned} \tag{5.9}$$

If the distance is an even number, the joint probability or correlation function at distance d' is related to that at distance $d'/2$. If the distance is odd, the joint probability at distance d' is related to the joint probabilities at two shorter distances $(d'+1)/2$ and $(d'-1)/2$. If $(d'+1)/2$ is again odd, the connection leads to four distances $(d'+3)/4$, $(d'+1)/4$, $(d'-1)/4$, and $(d'-3)/4$, and so on. In general, the shorter distances can either be even or odd, and all three transition matrices have to be used. In a special case, when $d'=2^n$ (n is an integer), only the matrix $\underline{T}(d'/2,d')$ determines the scaling exponent. In this case, by using the largest nontrivial eigenvalue λ_2 in Eq. (5.6), we have

$$c=-\frac{\log(1-p/2-q)}{\log(2)}\ . \tag{5.10}$$

If $p=q$, $c=-\log(1-3p/2)/\log(2)$. The exponent α for the power spectrum $P(f)\sim 1/f^{\alpha}$ is

$$\alpha=1-c=1+\frac{\log(1-p/2-q)}{\log(2)}\ . \tag{5.11}$$

If d' is not a power of 2, then the largest nontrivial eigenvalues from both $\underline{T}(d'/2,d')$ and $T((d'-1)/2,d')$ should be used.

Because it is easier to derive the transition probabilities for fixed-branching-ratio Lindenmayer systems, in Appendix C, I list the transition probabilities and the eigenvalues of the transition matrices for all two-symbol, fixed-branching-ratio-2 Lindenmayer systems. From these results, one can know immediately which rules can generate limiting sequences with $1/f$ spectra (when the largest nontrivial eigenvalue is close to 1 at some range of the parameter value) and which do not.

As a diversion, I will end this section by mentioning the local inhomogeneity for the sequences generated by Eq. (5.2). The last rewriting rule at Eq. (5.2) produces symbol 1 in the left leg but symbol 0 in the right leg. This distinction between the left and the right legs is a source of the local inhomogeneity. Suppose we ask the following question: go down the generating tree of rule (5.2) randomly; what is the probability of having symbol 1? Denote X_i^t as this probabilty at timer t and position i. Then, depending on which branch the random path goes through, there are two possibilities of how X_i^t changes:

$$X_i^t=\begin{cases}(1-q)X_{i/2}^t+p(1-X_{i/2}^t), & i \text{ even}\\ (1-q)X_{(i-1)/2}^t, & i \text{ odd}\ .\end{cases} \tag{5.12}$$

This random compositions of two linear maps is a one-dimensional case of the "iterative function systems."[21,22] Typically, there is no fixed point limit when $t\to\infty$; the X^t's are wandering around on a fractal set in the state space.[21,22] Figure 6 shows the "bifurcation diagram" for the mapping (5.12). The x axis is the parameter p (I assume $q=p$), and the y axis is the X^t with each point representing an orbital point of X^t (the site index i is suppressed).

If the trajectory $\{X^t\}$ consists of a fractal set, the fractal dimension D can be determined by the equation[21]

$$K_1^D+K_2^D=1\ , \tag{5.13}$$

where $K_1=1-p$ and $K_2=|1-2p|$ are the two slopes of the linear functions. At $p=\frac{1}{3}$, $D=1$ and at $p=\frac{1}{2}$, $D=0$. If the mutation rate is smaller than $\frac{1}{3}$, the dimension of the trajectory set remains at 1 (i.e., there is no fractal) as shown in Fig. 6.

Random compositions of two linear maps in the two-dimensional space are widely used to produce fractal patterns, or inversely, compress images with fractal components.[23] It is interesting that a special case of the "iterative function systems" finds an application in our expansion-modification sytems.

VI. CAN CELLULAR AUTOMATA GENERATE LIMITING SEQUENCES WITH $1/f$ SPECTRA?

Unlike Lindenmayer systems, cellular automata[17] are dynamical systems (or in certain sense, formal languages) which do not increase the sequence length. All generic cellular automata rules are "context sensitive," which means that the rewriting of the symbols not only depends

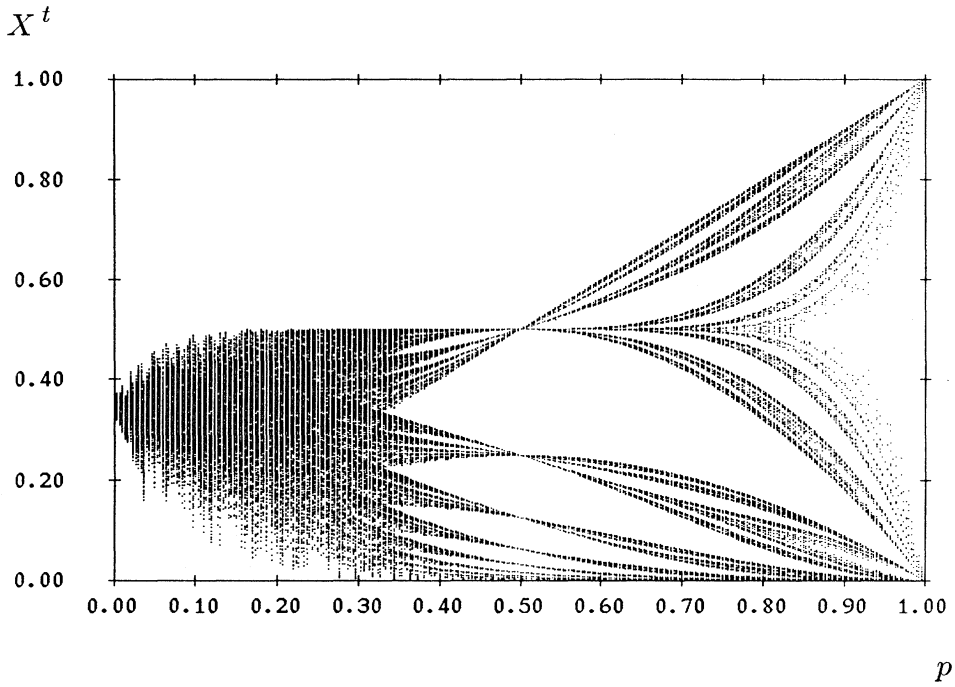

FIG. 6. Bifurcation diagram of the mapping (5.12) with two linear functions applied randomly (an example of the one-dimensional "iterative function systems"). The X axis is p and the y axis is X^t.

on the symbol value itself, but also depends on the symbol values at its neighboring sites. With this local interaction, it is possible that organizations of at least the neighborhood size can be produced from a random initial sequence. Some more interesting rules are capable of generating spatial correlations of much longer lengths than the neighborhood size. For example, the nearest-neighbor rule 110 [the rule maps the neighborhood configuration 000 to 0, 001 to 1, 010 to 1, 011 to 1, 100 to 0, 101 to 1, 110 to 1, and 111 to 0 (Ref. 24)] has limiting sequence containing periodic structures, and its periodicity 14 is much larger than the neighborhood size 3.

Long-range correlation alone cannot guarantee $1/f$ spectra though: $1/f$ spectra require that the long-range correlations scale with the short-range correlations, or the correlation function is a power-law function (with the exponent close to 0). It is not clear how to design a cellular automaton so that its limiting sequence has a $1/f$ spectrum.

Some earlier numerical studies on the lattice-map systems show that during the transient, there are signs of the temporal $1/f$ spectrum for the spatial-temporal spectrum $P(f,k)$ at some fixed k, but the spatial spectrum does not seem to be $1/f$.[25] Also, these $1/f$ spectra will not remain after the transients die out. A more recent study shows that the spatial $1/f$ spectra can exist during the transients in many cellular automata with soliton interactions.[26] Another study shows that the spatial block entropy of the limiting sequence of a nearest-neighbor rule 22 [the rule maps the neighborhood configuration 000 to 0, 001 to 1, 010 to 1, 011 to 0, 100 to 1, 101 to 0, 110 to 0, and 111 to 0 (Ref. 24)] does not increase linearly with the block length,[27] which implies that the correlation at longer lengths is more than what it should be for a random sequence (the difference is nevertheless very small).

The mechanism for producing $1/f$ spectra discussed in this paper cannot apply to cellular automata directly because the very source of the long-range correlations is by expanding the sequences, which cannot be accommodated by a cellular automaton. In order to mimic the expansion-modification by a probabilistic cellular automaton, consider the following rule with only left nearest-neighbor interactions ($x_{i-1}^t x_i^t \rightarrow x_i^{t+1}$:{probability}):

$$\begin{aligned} &00\rightarrow 0 \quad \{1/2+(1-p)/2\}, \\ &10\rightarrow 1 \quad \{1/2+p/2\}, \\ &01\rightarrow 0 \quad \{1/2+q/2\}, \\ &11\rightarrow 1 \quad \{1/2+(1-q)/2\}. \end{aligned} \tag{6.1}$$

The probabilities for other rewriting rules not listed above such as $00\rightarrow 1$ can simply be derived by the normalization condition: $P(00\rightarrow 1)=1-P(00\rightarrow 0)$.

The reason to write the rule in this way is that there is a tendency for x_{i-1} to "expand" (or "inject") its value to site x_i, and at the same time, there is a tendency for x_i to control itself by either mutating or maintaining its own value. There are different ways to assign weights to each competing factor. The probabilities used in Eq. (6.1) represent the following weight assignment: the "expansion" or the "injection"' from the left neighbor always

contributes a $\frac{1}{2}$ to the probability, and the "maintaining" of the site's own value contributes a $(1-p)/2$ or $(1-q)/2$ depending on whether the site value is 0 or 1.

When $p=q=0$, $00\rightarrow 0$ and $11\rightarrow 1$ always apply, so a block (a string of 0's or a string of 1's) will never be destroyed from inside. The basic structure of the spatial sequence is a long string of 0's followed by a long string of 1's, which in turn is followed by a string of 0's, and so on. Both the left and the right boundaries of a block move to the right with a probability 0.5. When the left boundary of a block moves faster than the right boundary the two can meet and consequently destroy the block. This destruction is irreversible, and the total number of blocks becomes smaller and smaller, or, the average block length becomes longer and longer. Figure 7(a) shows an example of this process (the time arrow points down), in which the final number of "black" blocks (and "white" blocks too) is only 3, much smaller than that at the earlier times.

Increasing the mutation rate a little bit, a block can be destroyed from inside with probability $p/2$. Figure 7(b) shows the spatial-temporal pattern at $p=0.1$. It already becomes extremely difficult to find a block with a long

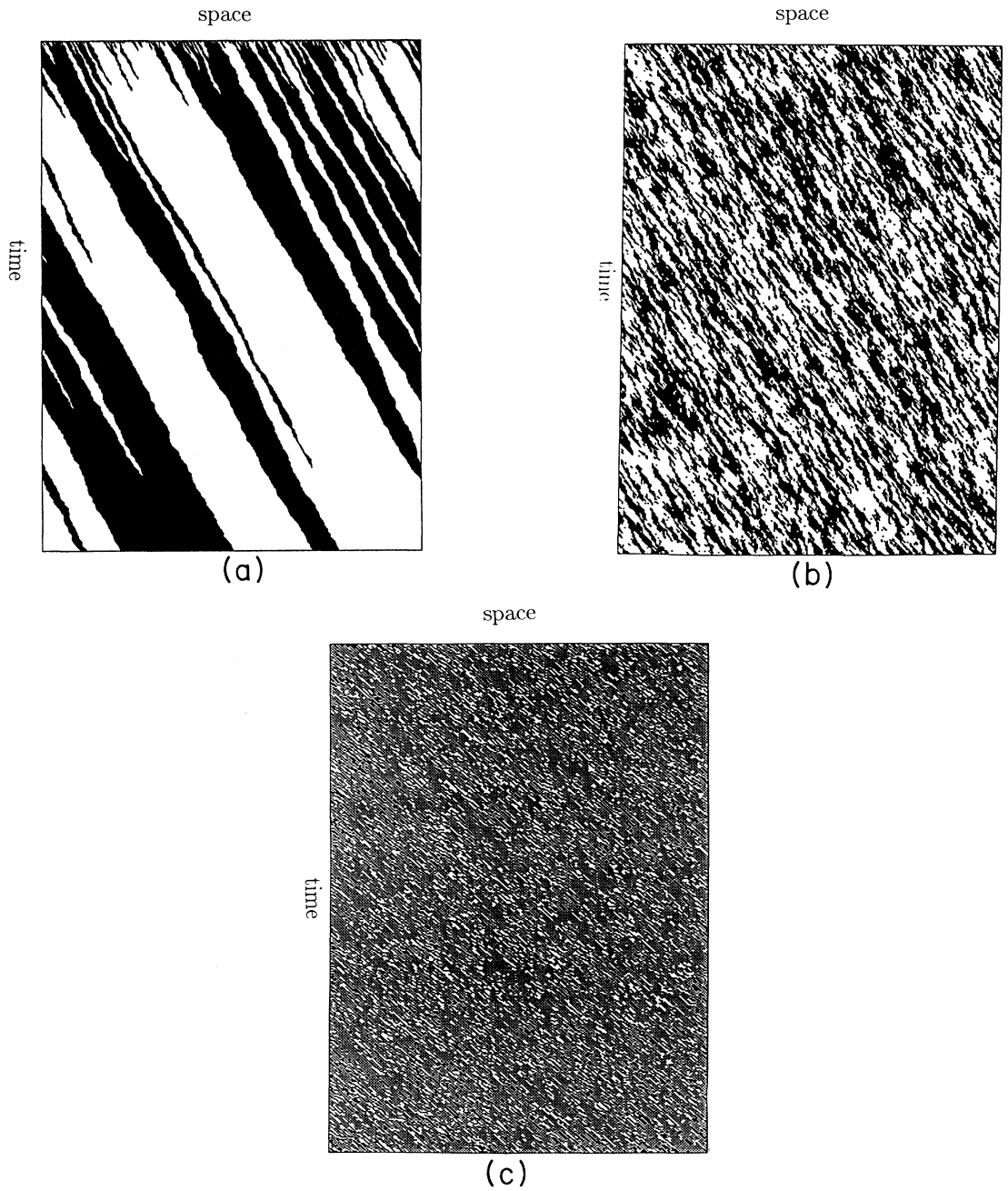

FIG. 7. Spatial-temporal patterns for the probabilistic cellular automaton rule (6.1) at $p=0$ (a); $p=0.1$ (b); and $p=0.9$ (c). The sequence length is 548 for (a), 274 for (b) and (c). The time step is 794 for (a), and 397 for (b) and (c).

block length. It is easy to imagine what the situation is like when the p is further increased. Figure 7(c) shows the spatial-temporal pattern at $p=0.9$.

As a measure of the correlations, the power spectra for the sequences generated at $p=0$ and 0.1 are calculated. Figure 8(a) shows the result at $p=0$. Although the limiting sequence for a finite lattice simulation might be all 0's or all 1's due to the irreversible destruction of blocks, the spatial spectra seem to have a $1/f^2$ form if we do not wait for the whole sequence to become a single block. This $1/f^2$ spectrum is similar to that for the sequences with randomly packed blocks as discussed in Appendix A. On the other hand, the spatial spectrum at $p=0.1$ shown in Fig. 8(b) has a small-frequency region with $1/f$ scaling. The lower-frequency spectrum is, however, flat, because of the lack of blocks with longer lengths.

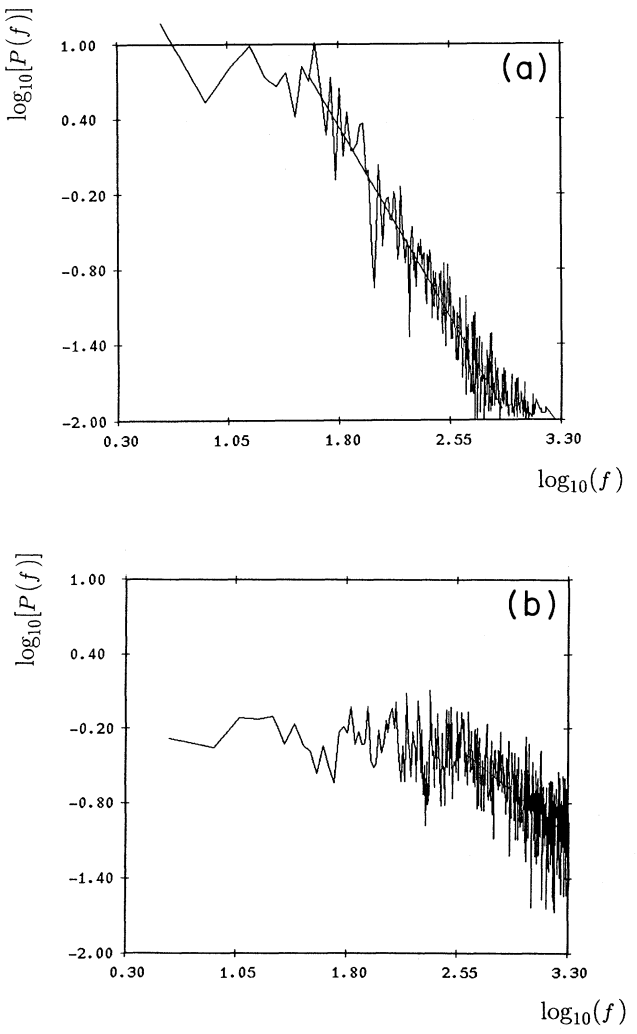

FIG. 8. The spatial power spectra of the sequences generated by the probabilistic cellular automaton rule (6.1). The sequence length is $2^{11}=2048$. (a) $p=0$, the straight line has the slope -2.04; (b) $p=0.1$, the short straight line has the slope -0.94.

VII. SUGGESTIONS OF THE APPLICATION

The expansion-modification systems can have applications to real-world systems if these systems contain both elements of creating and destroying the long-range correlations. Expanding, growing, or branching processes are all capable of generating long-range correlations, but without another balancing force to destroy the perfect order, the correlation would not decay at longer distances. Long-range correlation itself cannot guarantee scalings of the correlation as well as the $1/f$ spectrum.

Take electromagnetic showers in cosmic rays and hadron showers in higher energy experiments, for example; they are full of branching processes, such as the pair production $\gamma \rightarrow e^+e^-$ (γ for photon, e^- for electron, and e^+ for positron), and the radiation of photons $e^+ \rightarrow e^+\gamma$, $e^- \rightarrow e^-\gamma$. There are also processes involving the interaction among branches such as the annihilations $e^+e^- \rightarrow \gamma\gamma$. But processes similar to the "modification" are missing; for example, the electrons are not converted to a vacuum. As a result, it is unlikely that the spatial arrangement of the particles in the end of a shower will show anything like $1/f$ spectra.

In this section, I will discuss three expanding systems, with the first two having close connection with the expansion-modification system and the third one less similarities. They are as follows: (i) the evolution of the matter distribution in the Universe, (ii) the evolution of the prebiotic nucleotide sequences, and (iii) the generation of natural language texts. In the evolution of the Universe, the stretching of the distances between two mass points due to the expansion of the space coexists with the shrinking of the distances due to the gravitational attraction. In the evolution of the prebiotic nucleotide sequences, the replication tends to generate long-range correlations whereas the mutation tends to destroy them. The generation of a natural language text, though also an expanding process, does not seem to have the two competing factors. I included this example because I want to show a real case which is related to the Lindenmayer system whose largest nontrivial eigenvalue of the transition matrix is negative.

The existing data provide evidence that there are indeed nontrivial long-range correlations in the two examples mentioned above. The matter distribution (it is actually the galaxy-galaxy correlation that is measured) of the Universe is known to have a power-law correlation function.[28,29] For the nucleotide sequences, the correlation function varies greatly from sequence to sequence. Although most of the protein-coding sequences do not have long-range correlations, the noncoding sequences and the sequence segments between genes sometimes do (a brief report on this result is included in Ref. 30, and more studies are in progress).

A. Mass distribution of the Universe

The expansion-modification system is highly reminiscent of the evolution of the expanding Universe, in which, on the one hand, two points in space move further away due to the expansion, but on the other hand, any two mass points will attract each other due to the gravi-

tational force. Without gravitational attraction, the spatial distribution of the matter at time t will just be an expanded version of the initial configuration at time $t=0$, and no new structure is generated. Without the expansion of the space, a collection of mass points with only gravitational interaction will collapse—not a complicated configuration either (although the dynamics for each mass point may be nontrivial). At a certain conceptual level, the expansion-modification system addresses exactly this issue concerning two competing processes, although one has to translate the term "modification" to "gravitational clustering" with much more realistic modelings.

The basic quantity to characterize the spatial configuration of the Universe at any instant of time is the two-point mass-mass correlation function. Higher-order correlation functions such as the three-point correlation function are also used, especially when the two-point correlation function alone cannot distinguish obviously different configurations (the case with a typical non-Gaussian character). Currently, the mass-mass correlation function is not measured directly, due to the large amount of "dark matter" (the material that does not emit or absorb radiations[31,32]) in the Universe. What is measured is the galaxy-galaxy correlation function[28,29] or the cluster-cluster (cluster of galaxies) correlation function.[33] The assumption is that the three correlation functions are proportional to each other.[34,35] These correlation functions, for both galaxy-galaxy pairs and the cluster-cluster pairs, are observed to be the power-law function

$$\Gamma(d)\sim\frac{1}{d^{\beta}},\quad \beta\approx 1.8 \tag{7.1}$$

correct within certain length scales. It has been mentioned that at larger scales, the correlation function does not fit the same power-law function,[36] i.e., there is a cutoff of the scaling behavior. It is not clear whether it is due to the limitation of the observational data or due to the finite lifetime of the Universe so that the larger scale structures have not got a chance to develop. For the most recent studies on the larger scale structures of the universe, see Ref. 37.

There is a large number of papers discussing various models which lead to the power-law correlation function of the mass distribution. The consensus of opinion is that the large scale structure of the Universe is developed only from the gravitational instability (for the earlier review see Ref. 28, and for a recent one see, e.g., Ref. 38). With this basic assumption, the evolution of the Universe can be simply studied by putting a number of mass points in the expanding space and seeing how they evolve. It is similar to the computer simulations of the molecular dynamics except that here the space is expanding and the interaction between points is gravitational instead of the short-ranged van der Waals force.

Many such computer simulations are carried out[39–43] and, depending on initial conditions, the power-law correlation function with the exponent close to 1.8 can indeed by reproduced. The readers can go back to the original references to get more information on the range of length scales with the power-law behavior, the fitting exponents, the initial conditions used, and the evolution times, etc. Although the success in reproducing the two-point correlation cannot guarantee that other cosmological observations are also reproduced, it at least shows that, with certain initial conditions, gravitational systems in an expanding space can produce spatial configurations with power-law correlation functions. The limitations of the "success" include the following: (i) the correlation function at larger scales may deviate from the power-law function, and (ii) the correlation function itself may change with time, indicating a discrepancy between the transient configurations and the limiting configuration.

From the point of view of the expansion-modification system, the spatial power-law correlation function with a certain exponent results from a correct balance between the expansion of the space and the gravitational clustering. By tuning the "strength" of either the expansion or the clustering, the power-law correlation function is conceivable destroyed, either by way of changing the exponent, changing the range of the scaling, or changing the functional form. The "strength" of the expansion is measured by the expansion rate, which is determined by the Hubble constant. The strength of the gravitational clustering is measured by the exponent $\gamma=2$ in the Newtonian form of gravitation: $F(r)=Gm_1m_2/r^{\gamma}$, and the mass density of the Universe.

If we increase the value of γ from 2 to some very large value, the interaction among mass points is relatively short ranged, and it is unlikely that the attraction can provide an effective counterforce against the expansion. Of course, such parameter tuning is not realistic. As for the tuning of the expansion rate, because the expansion of the Universe is also a gravitational effect (according to the theory of general relativity), the Hubble constant is a function of the mass density. The larger the mass density, the larger the expansion rate. Perhaps the expansion and the clustering processes are indeed "balanced" after all, because of this fact.

It is interesting to note that there are other models, some of them purely geometrical, that can reproduce the power-law correlation function. One example[44] is the so-called *Voronoi foams.* One can start from a distribution of points, then draw lines (or planes, if the space is three-dimensional) perpendicular to the links between points, and these partition lines consist of a Voronoi foam. The points represent the centers of the low mass density regions (voids), and due to the gravitational instability these low-density regions become even less dense, and the mass flows to the partition lines. The mass can continue to flow along the partition lines into the vertices. It is claimed that by this process the power-law correlation function can be recovered.[44]

Another example is also geometrical in a sense: the formation of the cosmic strings.[45] There is evidence that the string loops evolve in a self-similar manner, and the resulting correlation function for the mass residing around the string loops is a power-law function. Even a cellular automaton with a threshold function is shown to exhibit power-law correlation functions with the exponent changing with the threshold value.[46] However,

the exponent 1.8 can only be recovered at a particular parameter value in this model.

B. Prebiotic nucleotide sequences

Evolution of prebiotic nucleotide sequences is another example in which two competing processes play an important role in determining the statistical properties of the sequences. On the one hand, the nucleotide sequences replicate. The recent developments in molecular biology have shown that RNA sequences have the ability to self-splice[47] and the ability to catalyze other reactions, including the replication. This discovery makes the scenario possible that in the periodic environment, the RNA sequences can replicate themselves even without the protein enzymes.[48] As more and more copies of the same sequence are created, there are chances that they can link together into longer sequences. In this way, long-range correlation can be generated.

On the other hand, the replication in the prebiotic environment was hardly perfect. All the mistakes during the replication—mutation, insertion and deletion—have the effect of reducing the long-range correlations in the linked sequence. The comparison with the expansion-modification system is now clear: if the prebiotic evolution contains only replication, the linked limiting sequence is periodic; if the mutation rate is too high, the limiting sequence is random. Only when the two processes are in an appropriate balance, can the nucleotide sequences show nontrivial long-range correlations, for example, the power-law decay of correlations.

The reason I discuss only the prebiotic evolution instead of the present-day evolution is because the latter is much more complicated and it does not always increase the sequence length. The sophistication of the modern replication machinery with the many types of biomolecules involved in each step of the process would make any realistic models be very complicated. The expansion-modification system looks too simple to be relevant. Also, because the sequence length is hardly increased in the present-day evolution, the crucial mechanism to generate the long-range correlation is missing. Because the longer-range correlation does not necessarily imply the fitter species, biological systems do not have to increase their current nucleotide sequence length more in order to survive, though they can still explore new possibilities by shuffling segments in the existing sequence, e.g., the crossover.

The main obstacle in verifying the predictions of any mathematical models of the prebiotic evolution is the lack of available prebiotic nucleotide sequences. As a large approximation, let me assume that the present-day nucleotide sequences have many statistical properties in common with the prebiotic ones. Then we can infer the statistical properties from the present-day DNA sequences.

The search for long-range correlation and power-law decay of the correlation in present-day DNA sequences is being carried out. A brief result is reported in Ref. 30. The main observation is that the statistical properties of a DNA sequence depend on whether it is a protein-coding or a noncoding segement in a gene, and whether it is a gene or a segment between genes. Without a large number of sequences in each category accessible, it is difficult to make a general conclusion on what the consensus statistical property is. The situation will certainly be better as more and more genomes are sequenced in the ongoing human genome project.[49]

The preliminary analysis of a few available sequences reveals the following results: (i) First of all, it seems that most of the protein-coding sequences (called *exons*) are random, in the sense that the nucleotide-nucleotide correlation decays to zero after only a few bases. The correlation can be measured either by the correlation function for a particular symbol, or by the mutual information function which averages over pairs of all symbol states.[50] Since three nucleotides code for one protein, there is a natural unit of three-base. As the correlation decays to zero, it can show an oscillation of a period of three. (ii) Second, the noncoding sequences (called *introns*) have longer correlation lengths then exons. Many introns have some two-unit segment being repeated many times. It is likely that these repetitions are responsible for the observed long correlation lengths. The exons do not have these exact repetitions, otherwise, the protein sequence consequently translated will repeat the same amino acids many times, a situation not needed in a catalytic reaction.

If a DNA sequence is not decomposed into genes and segments between genes, or exons and introns, what is the correlation function or mutual information function like? In this case, the segments between genes contribute more to the statistics (because of their longer lengths), and the larger scale structures, such as the distribution of genes, should also become important. It has been discovered that in eukaryotic DNA sequences, it is a rule rather than exception that there are many repetitive segments.[51–53] Sometimes, the repeated segments are next to each other (*tandem repeat*), other times they are separated (*interspersed repeat*). The length of the repeated segments varies greatly; some of them are short, others are as long as the whole genes. With this large number of possible variations, one might hope that one or a few DNA sequences can have the power-law correlation function or $1/f$ spectrum. Indeed, a partial $1/f$ spectrum has been observed in one intron sequence.[30] However, the $1/f$ spectrum may not appear in DNA sequences as often as the $1/f$ noise appears in, for example, the fluctuation of resistivity in conducting materials.[54]

C. Constituent sequences and letter sequences of natural language texts

The generation of a sentence in a natural language follows the rules of grammar, which can be approximated by certain context-free languages. In this generation process, a sentence breaks down to a few major constituents such as noun phrase and verb phrase; each constituent then is decomposed into subconstitutents, subconstituents into sub-subconstituents, and finally into words. The constituent composition of a sentence can be generated by the following grammar rules (S for sentence, N for noun phrase or noun, V for verb phrase or verb):

$$S \rightarrow NV ,$$
$$N \rightarrow S , \tag{7.2}$$
$$V \rightarrow VN .$$

Combining the first two rules, and including the cases of trivial rewriting $N \rightarrow N$ and $V \rightarrow V$, we have the more symmetrical representation:

$$N \rightarrow \begin{cases} NV \\ N , \end{cases}$$
$$V \rightarrow \begin{cases} VN \\ V . \end{cases} \tag{7.3}$$

If we ask the question of what the limiting sequence is if each rule in Eq. (7.3) is applied many times with a fixed probability, we go back to the same question discussed in the previous sections in this paper concerning the statistical properties of the limiting sequence generated by probabilistic Lindenmayer systems. In particular, comparing rule (7.3) with the rule (4.1), the first and the third rewritings are the same " complementary replication" because a symbol becomes two different symbols, but the second and the fourth rewritings in (7.3) maintain the symbol instead of modifying the symbol as in the case of rule (4.1). If the probability of applying the first and the third rewritings is large, the limiting sequence of rule (7.3) should be very similar to that of rule (4.1). As we know, the largest nontrivial eigenvalue of the transition matrix for (4.1) is negative, and the same should be for rule (7.3), so there is no $1/f$ spectrum expected to be present here either.

If the $1/f$ spectrum is missing from the limiting sequence of Eq. (7.3), what about the constituent sequence of a real natural language text? And what about a letter sequence of a real natural language text? We have the following arguments: (i) In a natural language text, the grammar rules are applied only a few number of times, whereas for rule (7.3), each rewriting is applied many, many times. Obviously, if the grammar rules are applied too many times, the sentence will contain subsentences, sub-subsentences, etc., and it is too complicated for a normal person to comprehend. (ii) The grammar rules of the natural language are only responsible for the structure of one sentence, but not responsible for the correlation between neighboring sentences. (iii) There are long-range correlations through the meaning of words in natural language texts. These correlations cannot be measured by correlation functions, mutual information functions, nor other statistical quantities. One can detect these correlations only when some translational device (e.g., a dictionary) is used. (iv) A letter sequence and the corresponding constituent sequence of a natural language text can have different statistical properties. It is because a constituent, such as N for noun, can have many ways to be transformed into a word, such as *table* or *dog*. Different words can contribute differently to the statistical property of the letter sequence.

Due to the above consideration, the statistical properties of the limiting sequence of rule (7.3) cannot be transformed simply to those of the constituent sequence as well as the letter sequence of a natural language text. But generally speaking, it is difficult for the long-range correlation to be present in a real natural language text, because the correlation is easily interrupted by the starting of a new sentence, and by the random realization of a word.

A numerical calculation of the mutual information function (similar to the correlation function, see Ref. 50) for letter sequences of English and German texts is carried out in Ref. 55. Indeed, no long-range correlation has been observed. If repetition is important in lifting the value of correlation at large distances, as in the case of intron DNA sequences and perhaps in music, it is missing from natural language texts. Perhaps the better analogy of natural language texts is the exon DNA sequence. If we do find a $1/f$ spectrum in a natural language text (perhaps in a poem?) maybe the repetition is too frequent for that text to be acceptable?

VIII. CONCLUSION

In this paper, it is proposed that when two competing processes—expansion and "modification"—coexist, at certain ranges of the parameter value which measure the relative strength of each process, the sequences or patterns generated by the dynamics have nontrivial long-range correlation (periodic sequences are examples of the *trivial* long-range correlation). In particular, the long-range correlation function can be a power-law function with an extremely small exponent, so the corresponding power spectrum is $1/f$.

There are many interesting aspects of the system which are not studied here, such as the power-law divergence of the block entropy (some preliminary results were obtained,[56] similar to the study in Ref. 27), the multifractal description of the block length distribution in the limiting sequence (similar to the study in Ref. 57), and the absence of the temporal $1/f$ spectrum following the temporal sequence of, say, the first site of the sequence (it provides a specific example that the spatial $1/f$ spectrum does not lead to the temporal $1/f$ spectrum). Although the many length and time scales phenomenon is not the result of a single mechanism, what is discussed in this paper seems to provide a rather general model which potentially may have many applications.

ACKNOWLEDGMENTS

There are many people I would like to thank for discussions and suggesting references: K. Kaneko, N. Packard, and S. Wolfram (on the original idea); W. K. Wootters (on the calculation of transition probabilities); J. Barnes, J. Scalo, R. van der Marel, D. Richstone, and R. di Stefano (on the matter distribution in the Universe); R. Fox, C. Burk, R. Bagley, and W. Fontana (on nucleotide sequences); M. de Sousa Vieira and M. Casdagli (on iterative functional systems). I thank the Aspen Center for Physics for providing an interactive environment. The financial support from the Recognizing Genes Workshop held at Aspen is also very much appreciated. This project was started at the Center for Complex Sys-

tems Research at University of Illinois, where the work was funded by National Science Foundation (NSF) Grant No. PHY-86-58062 and Office of Naval Research Grant No. N-00014-88-K-0293. The work at the Santa Fe Institute was funded by the MacArthur Foundation, NSF Grant No. PHY-87-14918, and U.S. Department of Energy Grant No. DE-FG05-88ER25054.

APPENDIX A: CORRELATION FUNCTION FOR SEQUENCES WITH RANDOMLY PACKED BLOCKS [SEE EQ. (2.3)]

This appendix derives the joint probability (and the correlation function) for sequences with randomly-packed "blocks" (all blocks have the same length m, "1-block" means a string of 1's, and "0-block" means a string of 0's):

$$\underbrace{00\cdots0}_{m}\ \underbrace{11\cdots1}_{m}\ \underbrace{11\cdots1}_{m}\ \underbrace{00\cdots0}_{m}\ \underbrace{11\cdots1}_{m}\ . \tag{A1}$$

To apply the result to the sequences generated by Eq. (2.1) at $p=0$ starting from a random sequence, simply replace m by 2^t, where t is the number of time steps.

The joint probability of two symbol 1's separated by distance d is represented by $P_{11}(d)$. The correlation function $\Gamma(d)$ is equal to $P_{11}(d)-P_1^2$ (P_1 is the density of symbol 1). For sequences with only two symbols, it has been shown that other joint probabilities such as $P_{10}(d)$ have the same functional form as that of $P_{11}(d)$. So once we know the $P_{11}(d)$, we know all other joint probabilities.[50]

When $d<m$, we count all the 1-1 pairs within 1-block's and between two neighboring 1-block's. Suppose the number of 1-block's is N_1, and the number of neighboring 1-block's is $N_{11}[1]$. If the periodic boundary condition is used, the last site in the sequence is considered as the left neighbor of the first site of the sequence. It is easy to show that

$$P_{11}(d)=\frac{N_1}{N}(m-d)+\frac{N_{11}[1]}{N}d \quad \text{if } 0\le d<m \tag{A2}$$

where N is the sequence length.

If $m\le d<2m$, we have to count the 1-1 pairs within two neighboring 1-block's, and in two "next-nearest-neighbor" 1-block's, i.e., two 1-blocks separated by another block (there are $N_{11}[2]$ of them). We have

$$P_{11}(d)=\frac{N_{11}[1]}{N}(2m-d)+\frac{N_{11}[2]}{N}(d-m) \quad \text{if } m\le d<2m\ . \tag{A3}$$

For arbitrary distances, suppose we know the number of two 1-block pairs with $i-1$ other blocks in-between (equal to $N_{11}[i]$); the general formula for the joint probability is

$$P_{11}(d)=\frac{N_{11}[i-1]}{N}(im-d)+\frac{N_{11}[i]}{N}[d-(i-1)m]$$
$$=\frac{iN_{11}[i-1]-(i-1)N_{11}[i]}{N}m-\frac{N_{11}[i-1]-N_{11}[i]}{N}d$$
$$\text{if } (i-1)m\le d<im\ . \tag{A4}$$

When all the $\{N_{11}[i]\}$'s are known, the $P_{11}(d)$ is known exactly, which is a piecewise linear function. If the case of random packing, $N_{11}[i]$ does not change with i very much. Depending on how $N_{11}[i]$ fluctuates with i, the correlation function can either increase linearly, decrease linearly, or stay at a constant. This piecewise linear correlation function leads to a $1/f^2$ power spectrum.

APPENDIX B: TRANSITION PROBABILITIES FOR THE SIMPLEST EXPANSION-MODIFICATION SYSTEM [SEE EQS. (3.6) AND (4.2)]

The transition probability $T(\alpha\beta d\to\alpha'\beta'd')$ is the probability of the symbol pair α-β separated by distance d leading to the symbol pair α'-β' separated by distance d' upon rewriting of a context-free Lindenmayer system. For the simplest expansion-modification system (2.1), the rewriting of one symbol can be classified into two types: doubling (the graph representing this type of rewriting is called a *scissors)* and mutation (the graph representing this rewriting is called a *bar)* (see Fig. 9). The relations between d and d' depend on whether one of the symbols or both symbols of α, β double or mutate.

We can further classify the rewriting of two symbols into six cases, as illustrated in Fig. 10. In cases A_1, A_2, and A_3, there are two scissors, but symbols α' and β' are either both on the two outer legs, or one on the outer leg and another on the inner leg, or both on the inner legs. The case A_2 is actually used twice: α' is on the inner leg and β' is on the outer leg, or the reverse. In cases B_1 and B_2, there is one scissors and one bar; again, both cases will be counted twice. In case C, there are two bars.

Suppose we know that symbol pair α-β leads to α'-β' by one of the six cases mentioned above; what is the probability that it can be accomplished, given that the distance between α and β is d and the distance between α' and β' is d'? If all the symbols in-between α and β double, d' is almost double d. On the other hand, if all the symbols mutate, d' should be the same as d. If the two distances d' and d are given, roughly speaking, $d'-d$ of the symbols double, whereas the remaining symbols (approximately $2d-d'$ of them) mutate.

Take case A_1 for example; the number of symbols in-between α and β that double is $d'-d-2$ (because α and β are already known to contribute to the increase of d' from d), and the number of symbols in-between α and β that mutate is $2d-d'+1$. Since the probability for a symbol to double is $1-p$ and that for a symbol to mutate is p, the overall probability for case A_1 is proportional to

$$\begin{bmatrix} d-1 \\ (d-1)-(d'-d-2) \end{bmatrix} p^{2d-d'+1}(1-p)^{d'-d-2}(1-p)^2 \ . \tag{B1}$$

The $(1-p)^2$ factor at the end of the line is the probability for the two end symbols α and β to double. Using a shorthand notation:

$$m \equiv (d-1)-(d'-d)=2d-d'+1 \tag{B2}$$

and the binomial coefficient

$$B_m(n,p) \equiv \begin{bmatrix} n \\ m \end{bmatrix} p^m(1-p)^{n-m} \ , \tag{B3}$$

and considering the normalization factor $(2-p)$, the probability for case A_1 is

$$p(A_1)=\frac{(1-p)^2}{2-p}B_{m+2}(d-1,p) \ . \tag{B4}$$

Similarly, the probabilities for other cases are

$$\begin{aligned} p(A_2)&=\frac{(1-p)^2}{2-p}B_{m+1}(d-1,p) \ , \\ p(A_3)&=\frac{(1-p)^2}{2-p}B_m(d-1,p) \ , \\ p(B_1)&=\frac{p(1-p)}{2-p}B_{m+1}(d-1,p) \ , \\ p(B_2)&=\frac{p(1-p)}{2-p}B_m(d-1,p) \ , \\ p(C)&=\frac{p^2}{2-p}B_m(d-1,p) \ . \end{aligned} \tag{B5}$$

The probabilities for these configurations can be easily transformed to the transition probabilities for a particular Lindenmayer system. For example, with the rule (2.1), since cases A_1, A_2, and A_3 keep the symbol states unchanged, we have

$$T(\alpha\beta d \to \alpha\beta d')=\sum_i p(A_i)$$
$$=\frac{(1-p)^2}{2-p}(B_{m+2}+2B_{m+1}+B_m) \ . \tag{B6}$$

The factor of 2 for the B_{m+1} term is because case A_2 contributes twice. Similarly,

$$T(\alpha\beta d \to \bar{\alpha}\beta d')=T(\alpha\beta d \to \alpha\bar{\beta}d')$$
$$=\sum_i p(B_i)=\frac{p(1-p)}{2-p}(B_{m+1}+B_m) \tag{B7}$$

and

$$T(\alpha\beta d \to \bar{\alpha}\bar{\beta}d')=p(C)=\frac{p^2}{2-p}B_m \ . \tag{B8}$$

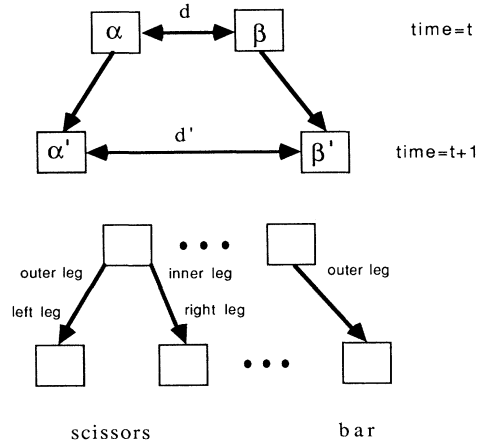

FIG. 9. Illustration of the notations and the names to be used in Appendix B.

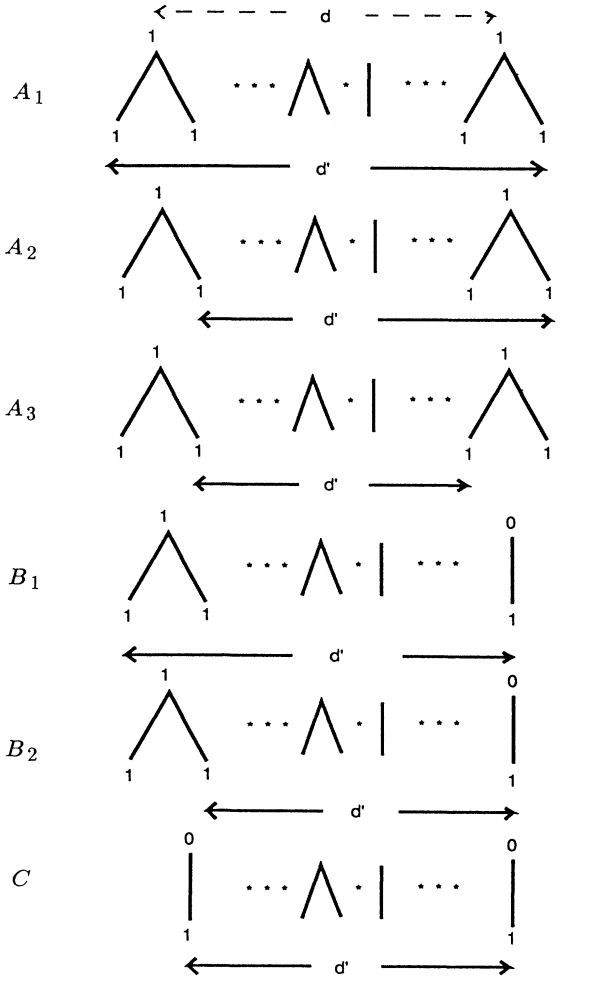

FIG. 10. Illustration of all the possible situations when "scissors" and "bar" are present. In the first three cases (A_1, A_2, and A_3), both ends have "scissors." In B_1 and B_2, there is one scissors and one bar. In C, there are two bars.

These results are used in Sec. III.

For rule (4.1), scissors can change the symbol, so $T(\alpha\beta d\rightarrow\alpha'\beta'd')$ take into account both A_i, B_i, and C. It is easy to show that the transition probabilities for this rule are

$$\begin{aligned}
&T(\alpha\beta d\rightarrow\alpha\beta d')=p(A_2)=\frac{(1-p)^2}{2-p}B_{m+1}\ ,\\
&T(\alpha\beta d\rightarrow\bar{\alpha}\beta d')=p(A_3)+p(B_2)=\frac{1-p}{2-p}B_m\ ,\\
&T(\alpha\beta d\rightarrow\alpha\bar{\beta}d')=p(A_1)+p(B_1)\\
&\qquad=\frac{(1-p)^2}{2-p}B_{m+2}+\frac{p(1-p)}{2-p}B_{m+1}\ ,\\
&T(\alpha\beta d\rightarrow\bar{\alpha}\bar{\beta}d')=p(A_2)+p(B_1)+p(B_2)+p(C)\\
&\qquad=\frac{1-p}{2-p}B_{m+1}+\frac{p}{2-p}B_m\ .
\end{aligned}\tag{B9}$$

APPENDIX C: TRANSITION PROBABILITIES AND THE EIGENVALUES OF THE TRANSITION MATRICES

In this appendix, I will list the transition probabilities $T(\alpha\beta d'/2\rightarrow\alpha'\beta'd')$ for all two-symbol Lindenmayer systems whose branching ratio is always 2 (see Sec. V). I will use the notation

$$a\rightarrow C_{2b_1+b_2,2c_1+c_2}[p]\tag{C1}$$

to represent the following rewriting rules for symbol a:

$$a\rightarrow b_1b_2\ \ \{1-p\},\quad a\rightarrow c_1c_2\ \ \{p\}\ ,\tag{C2}$$

with the parameter in curly brackets being the probability for that rewriting to be applied; and the notation

$$0\rightarrow C_{2b_1+b_2,2c_1+c_2}[p],$$

$$1\rightarrow C_{2b_1'+b_2',2c_1'+c_2'}[q]$$

(or sometimes simply $C_{2b_1+b_2,2c_1+c_2}C_{2b_1'+b_2',2c_1'+c_2'}$) to represent the Lindenmayer system:

$$\begin{aligned}
&0\rightarrow b_1b_2\ \ \{1-p\},\quad 0\rightarrow c_1c_2\ \ \{p\}\ ,\\
&1\rightarrow b_1'b_2'\ \ \{1-q\},\quad 1\rightarrow c_1'c_2'\ \ \{q\}\ .
\end{aligned}\tag{C3}$$

For example, Eq. (5.2) is $C_{02}[p]C_{30}[q]$ by this notation.

Equation (5.2) can also be represented by $C_{02}[p]C_{03}[1-q]$ by switching the two rewriting rules for symbol 1 and their probabilities. In general, we have $C_{ij}[p]=C_{ji}[1-p]$. The case of C_{ii} is simply a special example of $C_{ij}[p=0]$, and should not be counted as an independent case. With these considerations, the number of independent rewriting rules for one symbol is six. They are C_{01}, C_{02}, C_{03}, C_{12}, C_{13}, and C_{23}.

From these six independent rewriting rules for one symbol, one can construct 36 two-symbol Lindenmayer systems. But due to other equivalence relations, the number of independent Lindenmayer systems is further reduced. The first set of equivalence relations (represented by $\Longleftrightarrow$) is by switching symbols 0 and 1. We have

$$\begin{aligned}
&0\rightarrow C_{01}[p]\Longleftrightarrow 1\rightarrow C_{23}[1-p]\ ,\\
&0\rightarrow C_{02}[p]\Longleftrightarrow 1\rightarrow C_{13}[1-p]\ ,\\
&0\rightarrow C_{03}[p]\Longleftrightarrow 1\rightarrow C_{03}[1-p]\ ,\\
&0\rightarrow C_{12}[p]\Longleftrightarrow 1\rightarrow C_{12}[1-p]\ .
\end{aligned}\tag{C4}$$

Another set of transformations is the left-to-right transformation. The equivalence relations corresponding to this transformation are

$$\begin{aligned}
&0\rightarrow C_{01}[p]\Longleftrightarrow 0\rightarrow C_{02}[p]\ ,\\
&0\rightarrow C_{03}[p]\Longleftrightarrow 0\rightarrow C_{03}[p]\ ,\\
&0\rightarrow C_{12}[p]\Longleftrightarrow 0\rightarrow C_{12}[1-p]\ ,\\
&0\rightarrow C_{13}[p]\Longleftrightarrow 0\rightarrow C_{23}[p]\ .
\end{aligned}\tag{C5}$$

After using all the equivalence relations, the number of independent Lindenmayer systems in only 13.

The following results are derived by using the symbolic manipulation program MATHEMATICA.[58] Some explanations are as follows: (i) The transition probabilities from a symbol to another symbol through both the left and the right leg are represented by the array

$$\{\{T(0\rightarrow0),T(0\rightarrow1),T(1\rightarrow0),T(1\rightarrow1)\}_{\text{left}},\{T(0\rightarrow0),T(0\rightarrow1),T(1\rightarrow0),T(1\rightarrow1)\}_{\text{right}}\}\ .$$

(ii) The transition matrix can be constructed from the above array by [see Eq. (5.3)]

$$T_{2\alpha'+\beta',2\alpha+\beta}\equiv\frac{[T_{\text{left}}(\alpha\rightarrow\alpha')T_{\text{left}}(\beta\rightarrow\beta')+T_{\text{right}}(\alpha\rightarrow\alpha')T_{\text{right}}(\beta\rightarrow\beta')]}{2}\ .$$

These transition matrices are not listed here to save space. (iii) The eigenvalues of the matrix are

$$\{\lambda_i\}=\{\lambda_1,\lambda_2,\lambda_3,\lambda_4\} .$$

Note that the largest eigenvalue λ_1 is always equal to 1. For the 13 independent rules, we have the following.

(i) $0\rightarrow C_{01}[p]$, $1\rightarrow C_{01}[q]$:

$$\{T_{\text{left}},T_{\text{right}}\}=\{\{1,0,1,0\},\{1-p,p,1-q,q\}\} ,$$

$$\{\lambda_i\}=\left\{1,\frac{q-p}{2},\frac{q-p}{2},\frac{(q-p)^2}{2}\right\} .$$

(ii) $0\rightarrow C_{01}[p]$, $1\rightarrow C_{02}[q]$:

$$\{T_{\text{left}},T_{\text{right}}\}=\{\{1,0,1-q,q\},\{1-p,p,1,0\}\} ,$$

$$\{\lambda_i\}=\left\{1,\frac{q-p}{2},\frac{q-p}{2},\frac{p^2+q^2}{2}\right\} .$$

(iii) $0\rightarrow C_{01}[p]$, $1\rightarrow C_{03}[q]$:

$$\{T_{\text{left}},T_{\text{right}}\}=\{\{1,0,1-q,q\},\{1-p,p,1-q,q\}\} ,$$

$$\{\lambda_i\}=\left\{1,\frac{2q-p}{2},\frac{2q-p}{2},\frac{p^2-2pq+2q^2}{2}\right\} .$$

(iv) $0\rightarrow C_{01}[p]$, $1\rightarrow C_{12}[q]$:

$$\{T_{\text{left}},T_{\text{right}}\}=\{\{1,0,1-q,q\},\{1-p,p,q,1-q\}\} ,$$

$$\{\lambda_i\}=\left\{1,\frac{1-p}{2},\frac{1-p}{2},\frac{1-2p-2q+p^2+2pq+2q^2}{2}\right\} .$$

(v) $0\rightarrow C_{01}[p]$, $1\rightarrow C_{13}[q]$:

$$\{T_{\text{left}},T_{\text{right}}\}=\{\{1,0,1-q,q\},\{1-p,p,0,1\}\} ,$$

$$\{\lambda_i\}=\left\{1,\frac{1-p+q}{2},\frac{1-p+q}{2},\frac{1-2p+p^2+q^2}{2}\right\} .$$

(vi) $0\rightarrow C_{01}[p]$, $1\rightarrow C_{23}[q]$:

$$\{T_{\text{left}},T_{\text{right}}\}=\{\{1,0,0,1\},\{1-p,p,1-q,q\}\} ,$$

$$\{\lambda_i\}=\left\{1,\frac{1-p+q}{2},\frac{1-p+q}{2},\frac{1+p^2+q^2-2pq}{2}\right\} .$$

(vii) $0\rightarrow C_{03}[p]$, $1\rightarrow C_{01}[q]$:

$$\{T_{\text{left}},T_{\text{right}}\}=\{\{1-p,p,1,0\},\{1-p,p,1-q,q\}\} ,$$

$$\{\lambda_i\}=\left\{1,\frac{q-2p}{2},\frac{q-2p}{2},\frac{2p^2+q^2-2pq}{2}\right\} .$$

(viii) $0\rightarrow C_{03}[p]$, $1\rightarrow C_{03}[q]$:

$$\{T_{\text{left}},T_{\text{right}}\}=\{\{1-p,p,1-q,q\},\{1-p,p,1-q,q\}\} ,$$

$$\{\lambda_i\}=\{1,q-p,q-p,(q-p)^2\} .$$

(ix) $0\rightarrow C_{03}[p]$, $1\rightarrow C_{12}[q]$:

$$\{T_{\text{left}},T_{\text{right}}\}=\{\{1-p,p,1-q,q\},\{1-p,p,q,1-q\}\} ,$$

$$\{\lambda_i\}=\left\{1,\frac{1-2p}{2},\frac{1-2p}{2},\frac{1-2p-2q+2p^2+2q^2}{2}\right\} .$$

(x) $0\rightarrow C_{12}[p]$, $1\rightarrow C_{01}[q]$:

$$\{T_{\text{left}},T_{\text{right}}\}=\{\{1-p,p,1,0\},\{p,1-p,1-q,q\}\} ,$$

$$\{\lambda_i\}=\left\{1,\frac{q-1}{2},\frac{q-1}{2},\frac{1-2p-2q+2p^2+q^2+2pq}{2}\right\} .$$

(xi) $0\rightarrow C_{12}[p]$, $1\rightarrow C_{12}[q]$:

$$\{T_{\text{left}},T_{\text{right}}\}=\{\{1-p,p,1-q,q\},\{1-p,p,1-q,q\}\} ,$$

$$\{\lambda_i\}=\{1,0,0,(p-q)^2\} .$$

(xii) $0\rightarrow C_{13}[p]$, $1\rightarrow C_{01}[q]$:

$$\{T_{\text{left}},T_{\text{right}}\}=\{\{1-p,p,1,0\},\{0,1,1-q,q\}\} ,$$

$$\{\lambda_i\}=\left\{1,\frac{q-p-1}{2},\frac{q-p-1}{2},\frac{1-2q+p^2+q^2}{2}\right\} .$$

(xiii) $0\rightarrow C_{13}[p]$, $1\rightarrow C_{02}[q]$:

$$\{T_{\text{left}},T_{\text{right}}\}=\{\{1-p,p,1-q,q\},\{0,1,1,0\}\} ,$$

$$\{\lambda_i\}=\left\{1,\frac{q-p-1}{2},\frac{q-p-1}{2},\frac{1+p^2+q^2-2pq}{2}\right\} .$$

To check the result, take rule (5.2) for example, which is also $C_{02}[p]C_{03}[1-q]$, or $C_{01}[p]C_{03}[1-q]$. By substituting q with $1-q$, we get the largest nontrivial eigenvalue $\lambda_2=1-q-p/2$, the same as what we have derived in Eq. (5.6).

These results of eigenvalues of the transition matrices can immediately tell which Lindenmayer system at what parameter value can generate the limiting sequence with $1/f^{\alpha}$ spectra ($\alpha\approx 1$): it is when the largest nontrivial eigenvalue $\lambda\approx 1$, then the exponent for the correlation function $c=-\log(\lambda)/\log(2)\approx 0$, and the exponent for the spectrum $\alpha=1-c\approx 1$.

[1]B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1982).
[2]See, e.g., W. H. Press, Comments Astron. **7**, 103 (1978).
[3]B. Alder and T. Wainwright, Phys. Rev. Lett. **18**, 988 (1967).
[4]Y. Pomeau and P. Manneville, Commun. Math. Phys. **77**, 189 (1980).
[5]K. G. Wilson, Sci. Am. **241**, 158 (1979).
[6]T. Hatch and S. Choate, J. Franklin Inst. **207**, 369 (1929); T. Hatch, *ibid.* **215**, 27 (1933).
[7]J. Aitchison and J. Brown, *The Lognormal Distribution* (Cambridge University Press, London, 1963).
[8]H. Furukawa, Prog. Theor. Phys. **75**, 195 (1986).

[9]H. Furukawa, Phys. Rev. A **34**, 2315 (1986).
[10]W. Li, Europhys. Lett. **10**, 395 (1989).
[11]A. Lindenmayer, J. Theor. Bio. **18**, 280 (1968).
[12]See, e.g., J. E. Hopcroft and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation* (Addison-Wesley, Reading, MA, 1979).
[13]N. Packard, in *The Economy as an Evolving Complex System,* edited by P. W. Anderson, K. J. Arrow, and D. Pines (Addison-Wesley, Reading, MA, 1988).
[14]W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C* (Cambridge University Press, London, 1988).
[15]See, e.g., S. Karlin and H. M. Taylor, *A Second Course in Stochastic Processes* (Academic, New York, 1981); S. Karlin, *A First Course in Stochastic Process* (Academic, New York, 1968).
[16]W. Li, Complex Syst. **1**, 107 (1987).
[17]S. Wolfram, Rev. Mod. Phys. **55**, 601 (1979).
[18]T. C. Halsey, M. H. Jensen, L. P. Kadanoff, I. Procaccia, and B. Shraiman, Phys. Rev. A **33**, 1141 (1986).
[19]A. Thue, Nor. Vid. Selsk. Skr. **7**, 1 (1906).
[20]H. M. Morse, Trans. Soc. **22**, 84 (1921).
[21]M. F. Barnsley and S. Demko, R. Soc. London Ser. A **399**, 243 (1985).
[22]M. F. Barnsley, *Fractal Everywhere* (Academic, San Diego, 1988).
[23]M. F. Barnsley and A. D. Sloan, Byte, 215 (1988).
[24]*Theory and Application of Cellular Automata,* edited by S. Wolfram (World Scientific, Singapore, 1986).
[25]K. Kaneko, Phys. Lett. A **125**, 25 (1987).
[26]Y. Aizawa, I. Nishikawa, and K. Kaneko, Physica D **45**, 307 (1990).
[27]P. Grassberger, J. Stat. Phys. **45**, 27 (1986).
[28]P. J. E. Peebles, *The Large-Scale Structure of the Universe* (Princeton University Press, Princeton, 1980).
[29]H. Totsuji and T Kihara, Pub. Astron. Soc. Jpn. **21**, 221 (1969).
[30]W. Li, Santa Fe Institute Technical Report No. SFI-91-02 (1991).
[31]G. Blumenthal, S. Faber, J. Primack, and M. Rees, Nature, **311**, 517 (1984).
[32]George R. Blumenthal, in *The Santa Fe TASI-87 Proceedings of the 1987 Theoretical Advanced Study Institute in Elementary Particle Physics,* edited by R. Slansky and G. West (World Scientific, Teaneck, NJ, 1988), Vol. 2.
[33]N. Bahcall and R. Soneira, Astrophys. J. **270**, 20 (1983).
[34]J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, Astrophys. J. **304**, 15 (1986).
[35]N. Kaiser, Astrophys. J. **284**, L9 (1984).
[36]P. J. E. Peebles, Physica D **38**, 273 (1989).
[37]W. Saunders *et al.*, Nature **349**, 32 (1991).
[38]J. R. Bond, in *Frontiers in Physics—From Colliders to Cosmology,* Proceedings of the 1989 Lake Louise Winter Institute, edited by B. Cambell and F. Khanna (World Scientific, Singapore, 1990).
[39]M. Davis, G. Efstathiou, G. Frenk, and S. White, Astrophys. J. **292**, 371 (1985).
[40]S. White, C. Frenk, M. Davis, and G. Efstathiou, Astrophys. J. **313**, 505 (1987).
[41]G. Efstathiou, C. Frenk, S. White, and M. Davis, Mon. Not. R. Astron. Soc. **235**, 715 (1988).
[42]S. Maddox, G. Efstathiou, W. Sutherland, and J. Loverday, Mon. Not. R. Astron. Soc. **242**, 43r (1990).
[43]C. Park, Mon. Not. R. Astron. Soc. **242**, 59r (1990).
[44]V. Icke and R. van de Weygaert, Astron. Astrophys. **184**, 16 (1987); **213**, 1 (1989).
[45]N. Turok, Phys. Rev. Lett. **55**, 1801 (1985).
[46]T. Vicsek and A. S. Szalay, Phys. Rev. Lett. **58**, 2818 (1987).
[47]T. R. Cech, Sci. Am. **255**, 64 (1986); T. R. Cech and B. L. Bass, Annu. Rev. Biochem. **55**, 599 (1986).
[48]T. R. Cech, Proc. Natl. Acad. Sci. **83**, 4360 (1986).
[49]J. D. Watson, Science **248**, 44 (1990).
[50]W. Li, J. Stat. Phys. **60**, 823 (1990).
[51]R. Britten and D. Kohne, Sci. Am. **222**, 24 (1970); Science **161**, 529 (1968).
[52]E. Long and I. Dawid, Annu. Rev. Biochem. **49**, 727 (1980).
[53]W. Jelinek and C. Schmid, Annu. Rev. Biochem. **51**, 813 (1982).
[54]M. Weissman, Rev. Mod. Phys. **60**, 537 (1988).
[55]W. Li, Santa Fe Institute Technical Report No. SFI-89-09 (1989).
[56]T. Meyer (private communication).
[57]L. P. Kadanoff, S. R. Nagel, L. Wu, and S. Zhou, Phys. Rev. A **39**, 6524 (1989).
[58]S. Wolfram, *Mathematica: A System for Doing Mathematics by Computer* (Addison-Wesley, Reading, MA, 1988).