Sensitivity of ballistic deposition to pseudorandom number generators

Raissa M. D'Souza
Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Yaneer Bar-Yam
New England Complex Systems Institute, 17 Cedar St., Newton, Massachusetts 02159

Mehran Kardar
Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(Received 23 December 1997)

Ballistic deposition (BD) serves as a prototype for studies of dynamic scaling phenomena in nonequilibrium growth processes. In BD, particles are sequentially added to a growing surface at randomly selected positions. The model is typically investigated by computer simulations where randomness is implemented by pseudorandom number generators (PRNGs). The implicit assumption that PRNGs adequately represent true randomness is tested in this study via a statistical analysis of the width of the BD interface. We study the width of the interface over time scales orders of magnitude longer than the expected model relaxation time, yet much smaller than the period of the PRNG, and observe fluctuations which still appear to be correlated. Distinct dynamic behavior is observed for an implementation with a different PRNG, further indicating a strong coupling between the model and the PRNGs (even with PRNGs that pass extensive statistical tests). Thus we demonstrate a breakdown of basic sampling assumptions, and of the ergodic exploration of phase space. [S1063-651X(98)04405-5]

PACS number(s): 05.40.+j, 68.70.+w, 75.40.Gb, 05.70.Ln

I. INTRODUCTION

Dynamic scaling phenomena in stochastic nonequilibrium systems have attracted increasing attention in recent years [1]. In theoretical models of open systems, the external influences are usually represented by random noise. In computer simulations of the models pseudorandom numbers (PRNs) implement the stochastic process. The use of deterministic PRN algorithms necessarily introduces some degree of correlation in the produced sequence of PRNs [2]. While these correlations are probably irrelevant in most applications, they may in principle couple to the underlying dynamics of the simulated model, resulting in artificial behaviors.

A good example of open nonequilibrium behavior is provided by growth of aggregates through random addition of particles. The width of the resulting interface exhibits dynamic fluctuations, which have been studied extensively, analytically, and numerically [3,4]. A particularly simple model for this phenomenon is ballistic deposition (BD) [5]: Particles are randomly placed above an aggregate growing on a substrate, they descend along a straight vertical path until they encounter a site on the existing cluster and stick there. The random placements of subsequent particles represent a stochastic process.

We present results of a numerical study of BD in which the potential coupling to PRNs is examined via statistical tests on the width of the growing interface. A dynamic change in a conjectured steady-state regime is observed, signaling a breakdown of ergodicity. The breakdown is quantified by demonstration of violations of the basic sampling assumptions. Statistically relevant inconsistencies occur repeatedly in the data. First, fluctuations statistically inconsistent with the steady-state distribution are observed (even when the time scale of observation is orders of magnitude greater than the expected relaxation time of the model). Second, values of steady-state quantities averaged over different sets of initial seeds are in statistical disagreement. Third, average values of steady-state quantities, obtained by using two distinct pseudorandom number generators (PRNGs), are in statistical disagreement. These results lead to the conclusion that the observed dynamical fluctuations are not inherent to BD, but result from a coupling to the PRNG algorithms.

The manuscript is organized as follows. Section II provides the algorithmic details of the BD model and its implementation, with emphasis on the role played by the PRNs. In Sec. III probability distributions, sampling assumptions, and the construction of the statistical tests are discussed. The statistical inconsistencies of the numerical data with the basic sampling and ergodicity assumptions are presented in Sec. IV. Potential implications for the use of PRNs, particularly in the context of growth models, are discussed in the concluding Sec. V.

II. MODEL AND IMPLEMENTATION

Given that the focus of this work is on the unwanted coupling of PRNGs with the underlying model, it is necessary to provide details of the numerical simulation in greater depth than usual. Here we shall review the BD algorithm, and its numerical implementation with specific PRNGs.

A. Ballistic deposition

In the BD model of growth, free particles initiated at random positions above a one-dimensional substrate descend ballistically and stick upon first touching the surface of the

growing cluster. The substrate of length L, consists of discrete columns indexed by integer values x, with $1 \leqslant x \leqslant L$. The growth interface is defined by the maximum occupied site along each column, $h(x,t)$, where $h(x,t)$ also takes on discrete integer values. Starting from a flat interface, $h(x,t=0)=0$ for all x, the surface evolves by sequential addition of particles to randomly chosen columns. The index number of particles deposited is denoted by t', and the deposition time by $t=t'/L$. Each deposition event consists of choosing a column, $x(t')$, by a call to a PRNG, and updating the height in that column as follows:

$$h(x(t'),t'+1)=\max[h(x(t')-1,t'),$$

$$h(x(t'),t')+1,h(x(t')+1,t')]. \quad (1)$$

Thus the deposited particle occupies the highest empty site with one or more occupied nearest neighbor sites; this mimics the process of cluster aggregation. The stochastic process in this model is the random choosing of successive columns.

While the resulting aggregates are compact, their interface is rough, with fluctuations that are expected to be self-similar at all scales. The width of the growth interface $\xi_L(t)$ on average increases following a power law behavior until reaching a steady asymptotic value, the magnitude of which depends on the underlying substrate size L. A good measure of $\xi_L(t)$ is the variance of the surface heights, $\{h(x,t)\}$,

$$\xi_L^2(t)=\frac{1}{L}\sum_{x=1}^{L}[h(x,t)-\overline{h(t)}]^2, \quad (2)$$

where $\overline{h(t)}$ is the mean height of the surface at time t.

It was originally pointed out by Family and Vicsek [6] that the scaling forms for the growth and saturation of the width of the growing interface can be described by a dynamic scaling ansatz, similar to that applicable to critical systems. Kardar, Parisi, and Zhang (KPZ) [7] introduced an analytic theory describing the evolution of fluctuations on growing surfaces, which has been successfully applied to several growth models. One consequence of KPZ theory is that the steady-state behavior for the interface fluctuations in one dimension should resemble a random walk; i.e., $\xi_L(t \rightarrow \infty) \propto L^{1/2}$.

B. Algorithmic details

Substrates of lengths $L=127$, 255, 511, 1023, 2047, and 10 007 are considered. At each update, a PRN is generated corresponding to a column along the substrate. A particle is added to that column at a height described by Eq. (1). Periodic boundary conditions are applied. The only subtlety is in mapping the PRN uniformly to a value between 0 and $L-1$, which is achieved as follows: The least significant bits of the PRN are shifted off, leaving a number between 0 and 2^n-1, with n chosen such that 2^n is the integer closest to, but greater than L. The PRN is rejected if it falls in the interval between L and 2^n-1. Variants on this scheme were tested (including use of all the bits of the PRN), yet similar results were obtained. The scheme described above was chosen because the algorithms that generate subsequent PRNs involve algebraic operations which cause carries from lower order bits to higher order bits; thus the higher order bits are influenced by two sources (the algebraic operation and the carries) and are expected to be less correlated.

As the system evolves, the variance of the surface heights, ξ_L^2, is calculated at selected times. Statistical errors are considered [8], and the associated standard error σ_{ξ^2} is also recorded. The values of the surface heights (contained in a one-dimensional array of length L) are the only essential data for these calculations.

The onset of the asymptotic regime is estimated by the conservative criterion of a ''relaxation'' time $\tau \sim 10L^{z'}$, with $z'=1.6$. This estimate is conservative since the exponent employed exceeds the expected value of the dynamic scaling exponent, $z=3/2$, and this time is well beyond a qualitative judgment of the time required for saturation of ξ_L^2. Accounting for the measured growth rate, this relaxation time corresponds to an average surface height, $\overline{h(\tau)}=20L^{z'}$. Exploring the asymptotic regime requires extensive computer power for the larger lengths investigated. All simulations were implemented on desktop workstations, with the shortest length systems requiring a few hours of run time, the longest requiring on the order of five days.

C. Randomness and PRNGs

PRNGs are algorithms for deterministically generating a string of bits, resembling a completely uncorrelated, and hence ''random'' string. Knowing the past and present values should give no information as to future outcomes of a truly random variable. Hence ''deterministic randomness'' is inherently unattainable. PRNGs are at best a practical substitute, and should be generally tested for the absence of undesired correlations. While two-point correlations can be readily examined, there are a multitude of other subtle effects that are not in practice possible to measure. When considering which tests for correlations to conduct, it is advisable to include both the standard statistical tests, as well as physically motivated ones directly related to the particular model being implemented (see the example in the final paragraph of this section). It is also necessary to verify that any observed dynamic behavior is inherent in the simulated model, and not artificially introduced to the system by the PRNGs. Several physical models have been shown to couple to correlations in PRNGs [9–11]; we shall provide evidence that BD also belongs to this category.

Preliminary simulations with a simple PRNG resulted in various anomalies which will not be discussed in detail here, but one example is the occurrence of repeated patterns of surface configurations. Having identified the PRNG as the likely culprit, we decided to use more sophisticated PRNGs. Several were tested, and two were selected: ''random(),'' a feedback shift register, employing a primitive trinomial of degree 64 [12] (a ''C-library'' subroutine call), and ''ran2()'' which combines pseudorandomness produced by two distinct multiplicative congruential generators, and has been shown to reduce certain serial correlations inherent to each generator separately [13].

Extensive tests for correlations in the PRN sequences were conducted, with emphasis on tests directly relevant to the BD growth algorithm. One such test is for any bias in the next relevant call. If the next growth site in a neighborhood is biased to the left- or right-neighboring column of the last

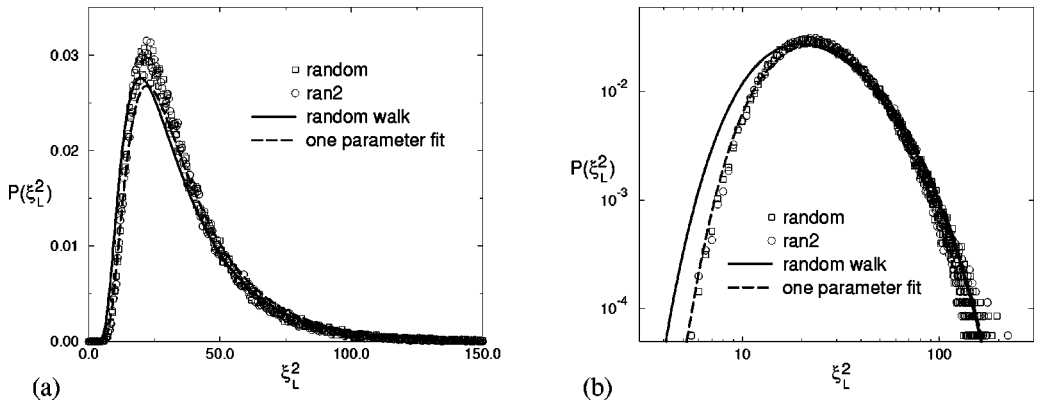

FIG. 1. The empirical steady-state distribution function $P(\xi_L^2)$ obtained for both PRNGs. The solid line corresponds to the theoretical distribution function for a random walk ($\alpha=1/2$), the dashed line to a one-parameter fit with $\alpha=0.45$, where α is the roughness exponent, as plotted on (a) a linear scale, (b) a log-log scale.

added particle, a thinner growth interface results. If it is biased towards the same column, a wider interface is generated. No discrepancies were found with any of the tests, a brief summary of which is included here. (1) Spatial and temporal Fourier transforms of the sequence of PRNs were consistent with white noise power spectrums, suggesting no two-point correlations. (2) Measurement of the number of calls to each column, and waiting times between subsequent calls to the same column, were consistent with Poisson statistics. (3) No spatial bias for the subsequent calls, as discussed above, was detected. (4) No bias was detected when the space was partitioned into sublattices. (5) Autocorrelation functions for natural surface height observables were decaying simple exponentials. The reader is referred to a previous manuscript for details [14].

III. PROBABILITY DISTRIBUTIONS

We first obtain the unique steady-state probability distribution for $\xi_L^2(t>\tau)$. Once this distribution is known, each independent measurement of $\xi_L^2(t>\tau)$ can be considered an independent, identically distributed (IID) random variable, drawn with the associated probability. Using only this general assumption of unbiased sampling, we construct statistical tests which show that implementations which are identical, except for use of different PRNGs, result in different values for average quantities of the growth interface; implementations with the PRNG studied most extensively in this work lack steady-state behavior; data obtained by averaging over several independent implementations are inconsistent with the underlying distribution.

The tests focus on the width of the growth interface in the steady-state regime $\xi_L^2(t>\tau)$. We shall discard the time argument in favor of a compact notation, and henceforth denote this variable by ξ_L^2. When it is necessary to deal with shorter times, the explicit time argument is included. The discussion also focuses on the $L=127$ system size. We were able to explore the asymptotic regime for orders of magnitude beyond the conjectured model relaxation time only for this shortest length investigated due to practical limits on computational resources. For the $L=127$ system, adequate statistics could be obtained for times as large as $t=1000\tau$.

A. The steady-state distribution

There is a unique, steady-state distribution, for the overall width of the growth interface, $P(\xi_L^2)$. It is shown in Fig. 1, as determined by sampling 350 realizations with different initial seeds for each PRNG; each realization was evolved to the asymptotic regime and ξ_L^2 was measured. Each realization was further evolved for ten autocorrelation times, and ξ_L^2 was measured again. This latter step was repeated 200 times. Hence the histogram shown in Fig. 1 was constructed with 7×10^4 data points for each PRNG. The data obtained by both PRNGs converge to the same empirical distribution, to well within statistical error. We will denote the average value of this distribution by μ, and its standard deviation by σ. For random() the values obtained are $\mu=35.87$ and $\sigma=20.85$. For ran2() the values obtained are $\mu=35.74$ and $\sigma=20.58$. Each independent observation of ξ^2, in the asymptotic regime, should be an IID random variable sampled from this distribution.

Before proceeding to the statistical tests, we briefly compare this distribution to previous ones obtained for growth models. The complete distribution function naturally contains much more information about the system than just the average value of ξ_L^2. The KPZ equation, as well as other exactly solvable models in one dimension [15,16], give rise to steady-state distributions which are identical to those of a random walk. The random walk distribution is expected to describe the BD model as well, assuming that it falls in the KPZ universality class. The theoretical distribution for the overall width of a random walk, with periodic boundary conditions, was calculated recently [17] and is shown in Fig. 1, overlaying the empirical BD distribution. There is a slight, but systematic disagreement between the BD histograms and the theoretical distribution for the random walk. The data for numerical implementations of the KPZ equation (and for solid-on-solid growth models) have been successfully fitted to the random walk distribution [17,18], confirming that this distribution does indeed describe systems in the KPZ universality class.

A *better* fit is obtained by following a phenomenological approach, introduced by Rácz and Plischke [18]. The width of the growth interface in the steady state can be obtained from the structure factor $S(k)=\langle\hat{h}(k)\hat{h}(-k)\rangle$, as ξ_L^2

$=b^{-1}\Sigma_{k\neq 0}S(k)$, where $\hat{h}(k)$ is the Fourier transform of $h(x)$ (and hence $k=2\pi m/L$, where m is an integer), and b is a constant with dimensions of inverse width squared. Assuming that $\{h(x)\}$ is continuous, and Gaussian distributed with a kernel given by $S(k)$, the probability distribution for the width is calculated as [18]

$$P(\xi_L^2)=\int_{-i\infty}^{i\infty}\frac{d\lambda}{2\pi i}e^{\lambda\xi^2}\prod_{k\neq 0}\frac{bS^{-1}(k)}{\lambda+bS^{-1}(k)}. \tag{3}$$

For self-similar fluctuations, the power spectrum behaves as $S(k)\propto|k|^{-\gamma}$, where γ is related to the standard roughness exponent α by $\gamma=2\alpha+d$, where d is the dimension of the substrate.

Figure 1 shows the phenomenological distribution function for the parameter value of $\alpha=0.45$, alongside the random walk distribution function ($\alpha=0.5$), and the numerical data. Figure 1(a) shows these functions plotted on a linear scale, Fig. 1(b) on a log-log scale. The phenomenological distribution shown is obtained by summing the residues of the first 45 poles in the contour integral of Eq. (3). The sum of the residues for the first 18 poles converges to the identical distribution throughout the regime considered, indicating that the result is essentially exact. The phenomenological distribution captures certain aspects of the numerical data with more fidelity than the random walk distribution. The value of $\alpha=0.45$ is smaller than the KPZ prediction of $\alpha=1/2$, but is consistent with previous values of the roughness exponent reported in numerical studies of BD [4,6,14,15].

In regard to the statistical tests discussed in the remainder of this manuscript, the relevant result presented in this section is that both PRNGs converge to the same statistical distribution. The discrepancies with random walk behavior will not be further considered in this manuscript.

B. Distribution of averages

We are concerned with the consistency between independent measurements, and introduce the statistical tests that follow in order to test this (or more accurately to evaluate the lack of consistency). The tests assume approximately Gaussian distributed variables. Although a single measurement of the width of the interface (Fig. 1) is not a Gaussian variable, the probability distribution of the average width over many realizations is Gaussian (via the central limit theorem [19]). In simulations one typically considers average quantities, which also necessitates knowing the probability distribution of the average. Our simulations are of N independent realizations, so the relevant distribution is that of the width of the interface averaged over N independent samples, $\overline{\xi_L^2}=\Sigma_{i=1}^{N}\xi_L^2/N$.

We can construct the actual distribution function for the average over N IID samples, $P(\overline{\xi_L^2})$, from its Fourier transform denoted by $\widetilde{P}_N(k)$ (and usually referred to as the characteristic function). The characteristic function of the average is related to the characteristic function of the individual samples $[\widetilde{P}(k)]$ by $\widetilde{P}_N(k)=[\widetilde{P}(k/N)]^N$. The characteristic function is the generator of the cumulants of the distribution, and the nth cumulant of the distribution $P(\overline{\xi_L^2})$ is related to the nth cumulant of the distribution $P(\xi_L^2)$ simply by $\langle(\overline{\xi_L^2})^n\rangle_c=(1/N^{n-1})\langle(\xi_L^2)^n\rangle_c$.

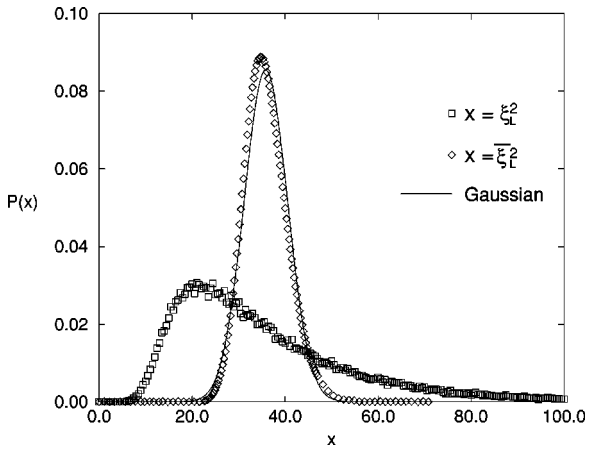

FIG. 2. The distribution function for the individual samples, $P(\xi_L^2)$, the distribution for the average over $N=20$ samples, $P(\overline{\xi_L^2})$, and a Gaussian approximation to the distribution for the average.

Dealing explicitly with the first two cumulants of the probability distribution of the average,

$$\langle\overline{\xi_L^2}\rangle=\left\langle\sum_{i=1}^{N}\frac{1}{N}\xi_{L,i}^2\right\rangle=\sum_{i=1}^{N}\frac{1}{N}\langle\xi_{L,i}^2\rangle=\mu, \tag{4}$$

$$\langle(\overline{\xi_L^2})^2\rangle_c=\sum_{i=1}^{N}\left\langle\left(\frac{1}{N}\xi_{L,i}^2\right)^2\right\rangle_c=\frac{1}{N^2}\sum_{i=1}^{N}\langle(\xi_{L,i}^2)^2\rangle_c=\frac{\sigma^2}{N}\equiv\nu^2. \tag{5}$$

Thus each independent observation of $\overline{\xi_L^2}$ is a random variable drawn from a distribution with mean μ and variance $\sigma^2/N\equiv\nu^2$.

As N increases in value, higher order cumulants go to zero, and the distribution approaches a Gaussian (as required by the central limit theorem [19]). The function $P(\overline{\xi_L^2})$ is shown in Fig. 2, along with the numerical data for $P(\xi_L^2)$ (previously shown in Fig. 1), and a Gaussian distribution with mean μ and variance ν^2, for $N=20$.

C. The χ^2 distribution

Each independent realization of $\overline{\xi_L^2}$ is approximately Gaussian distributed about the mean μ of the empirical distribution, with a variance ν^2 (see Fig. 2). Hence the normalized difference $(\overline{\xi_L^2}-\mu)/\nu$ should be a random variable sampled from a unit normal distribution (i.e., a Gaussian distribution with mean of zero, and unit variance). The sum of squares of M independently distributed unit normal random variables, denoted by $\chi_{\rm SS}^2$,

$$\chi_{\rm SS}^2=\sum_{i=1}^{M}\frac{(\overline{\xi_{L,i}^2}-\mu)^2}{\nu^2}, \tag{6}$$

follows a χ^2 distribution with M degrees of freedom [19]. If the $\chi_{\rm SS}^2$ statistic is sufficiently large, it is unlikely that all values in the sum are approximately unit normal distributed. The χ^2 test quantifies how unlikely; the test determines the

probability that a number of value χ^2_{SS} or greater is drawn from a χ^2 distribution with M degrees of freedom. We denote this test as the χ^2_{SS} test and use it to determine the probability of the hypothesis that all the values of $\overline{\xi_L^2}$ in the conjectured steady-state regime were sampled independently from the same underlying distribution (shown in Fig. 2).

When performing a simulation one uses the values of the average and standard deviation obtained in the simulation as estimates of the average and standard deviation of the distribution function. In order to use the sample average $\langle \overline{\xi_L^2} \rangle_M = (\Sigma_{i=1}^M \overline{\xi_{L,i}^2})/M$ in place of μ, and the weighted variance $s^2/(N-1) = \Sigma_{i=1}^N (\xi_{L,i}^2 - \overline{\xi_L^2})^2/[N(N-1)]$ in place of $\sigma^2/N = \nu^2$, we refer to the t test. Note that these values obtained from our simulations should be unbiased estimators of μ and ν^2, respectively [19], and represent more accurately the error bars obtained.

D. The t distribution

''Student'' [20] first discussed the error introduced by estimating σ^2 with the sample standard deviation $s^2 = \Sigma_{i=1}^N (\xi_{L,i}^2 - \overline{\xi_L^2})^2/N$, and suggested the t test as an alternative statistical test [20]. Note that $s^2/(\sigma^2/N) = s^2/\nu^2$ follows a χ^2 distribution with $N-1$ degrees of freedom (the degrees of freedom are reduced by 1 as there is one constraint on the random variables: the mean value is equal to $\overline{\xi_L^2}$). Following ''Student'' we can construct a statistic from the ratio of a unit normal distributed random variable $\mathcal{Z}$ to an independently χ^2 distributed random variable f_{N-1}^2 with $N-1$ degrees of freedom:

$$T_{N-1} = \frac{\mathcal{Z}}{\sqrt{f_{N-1}^2/(N-1)}}.$$

We label this variable T_{N-1}, as its probability density should follow a t distribution with $N-1$ degrees of freedom [19]. The unit normal distributed random variable we are interested in is $\mathcal{Z} = (\overline{\xi_L^2} - \mu)/\nu$. The χ^2 distributed random variable is $f_{N-1}^2 = s^2/\nu^2$. Thus the corresponding t statistic is

$$T_{N-1} = \frac{(\overline{\xi_L^2} - \mu)/\nu}{\sqrt{s^2/[(\nu^2)(N-1)]}} = \frac{(\overline{\xi_L^2} - \mu)}{\sqrt{s^2/(N-1)}}. \tag{7}$$

We likewise define a second t statistic, useful for comparing two independent data sets. Consider two independent sets of M IID samples drawn from a probability distribution with mean μ and variance ν^2 (i.e., two independent sets of M IID realizations of $\overline{\xi_L^2}$). There are thus two independent measurements of the average value over the M IID samples, denoted by $\langle \overline{\xi_L^2} \rangle_{M,i}$, for $i = 1,2$ (as defined earlier, but note the additional index i, used to designate the data set). Likewise there are two independent measurements of the variance over M samples, denoted by $\mathcal{S}_i^2 = \Sigma_{j=1}^M (\overline{\xi_{L,i,j}^2} - \langle \overline{\xi_L^2} \rangle_{M,i})^2/M$, for $i = 1,2$. The difference between two independent observations of average values is a random variable which converges to a Gaussian distribution with mean of zero and variance $2\nu^2/M$, in the limit of large M. Hence an approximately unit normal distributed random variable is $\mathcal{Z} = (\langle \overline{\xi_L^2} \rangle_{M,1} - \langle \overline{\xi_L^2} \rangle_{M,2})/\sqrt{2\nu^2/M}$. An independent χ^2 distributed random variable [with $2(M-1)$ degrees of freedom] is $f_{2(M-1)}^2 = \mathcal{S}_1^2/(\nu^2/M) + \mathcal{S}_2^2/(\nu^2/M)$. The t statistic to compare two independent data sets is thus

$$T_{2(M-1)} = \frac{(\langle \overline{\xi_L^2} \rangle_{M,1} - \langle \overline{\xi_L^2} \rangle_{M,2})}{\sqrt{(\mathcal{S}_1^2 + \mathcal{S}_2^2)[1/(M-1)]}}. \tag{8}$$

If the t statistic is sufficiently large, it is unlikely that the ratio of the sample mean to the sample variance is an accurate estimator of the ratio of the theoretical mean to the theoretical variance. The ''t test'' measures the probability that a value the size of the t statistic or greater is drawn from the appropriate t distribution [19]. In summary, the T_{N-1} test using the T_{N-1} statistic measures the level of validity for the hypothesis that each sample in a set was drawn from the same underlying distribution. The $T_{2(M-1)}$ test using the $T_{2(M-1)}$ statistic measures the level of validity for the hypothesis that samples in two distinct sets were drawn independently from the same underlying distribution. This latter test compares average values generated by the two distinct sets.

E. Distribution of extrema

The final class of statistical tests employed deals with extreme values found in sets of IID random variables. The extrema found in our data appear to present the largest deviation from the theoretical distribution. Of course, when dealing with extrema there is a selection effect, for which one has to correct. The probability of obtaining a particular value for the minimum in a set of M IID random variables can be readily calculated as [21]

$$p(x_{\min} = x) = M p(\overline{\xi_L^2} = x) \left(\int_x^\infty p(\overline{\xi_L^2}) d\overline{\xi_L^2} \right)^{M-1}. \tag{9}$$

Likewise the probability for obtaining a particular value for the maximum in a set of M IID random variables can be readily calculated as [21]

$$p(x_{\max} = x) = M p(\overline{\xi_L^2} = x) \left(\int_{-\infty}^x p(\overline{\xi_L^2}) d\overline{\xi_L^2} \right)^{M-1}. \tag{10}$$

The original probability distribution for the average width, $P(\overline{\xi_L^2})$, along with the distributions for the minimum and the the maximum in a set of $M = 10$ independent observations, are shown in Fig. 3.

IV. RESULTS OF STATISTICAL TESTS

We applied the statistical tests outlined above to the data obtained from our simulations. Results for the χ^2 test are reasonable, however, the results for the t test and for the extremal values are highly unreasonable. Thus specific simulations may give highly anomalous results, inconsistent with the theoretical distribution.

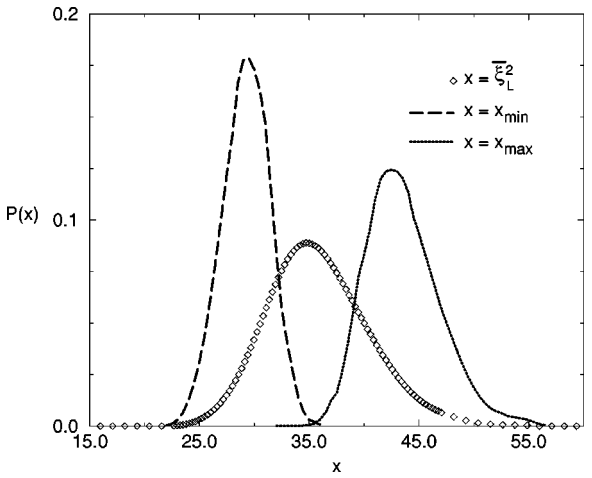

FIG. 3. The probability distribution for the average width, $P(\overline{\xi_L^2})$, along with the distribution for the minimum and the maximum in a set of $M=10$ independent observations of $\overline{\xi_L^2}$.

A. Comparison of data at different times (in the asymptotic regime)

In the asymptotic regime, the width of the interface should saturate to a steady-state value. The average over 20 independent samples, $\overline{\xi_L^2}$, for an implementation with random( ), is recorded at selected subsequent times, and a plot of these data is shown in Fig. 4. All of the values shown should be equal within statistical error, but there are large differences. In fact the greatest difference between two values is over four standard errors in magnitude. To quantify the significance of this difference, we apply the statistical tests discussed in Sec. III, to this data set.

The $\chi^2_{\rm SS}$ statistic, defined in Eq. (6), is calculated, but is not large enough to be significant: the probability that the null hypothesis of steady-state behavior is valid is 24%, as

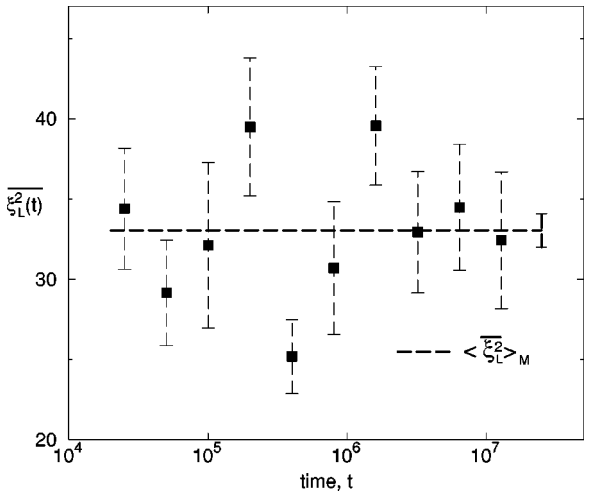

FIG. 4. The average width of the growth interface, $\overline{\xi_L^2}$, as obtained for 20 independent samples using random( ), shown at selected subsequent times in the asymptotic regime. The horizontal line corresponds to the average value $\langle\overline{\xi_L^2}\rangle_M$ of the $M=10$ data points in this figure, with the associated error bar plotted at the extreme end of the line. Note that the expected relaxation time is less than $\tau=2\times10^4$, and that the logarithmic scale spans roughly $10^3\tau$.

determined by the $\chi^2_{\rm SS}$ test. We denote this as a pass at the 24% confidence level (CL), or likewise a fail at the 76% CL. In other words an event with this magnitude is expected to occur one in four times, therefore a pass at the 24% CL is not an unreasonable result.

The T_{N-1} test for steady-state behavior, however, does manifest discrepancies. Eight of the ten points pass the T_{N-1} test above the 20% CL. But two points fail at the 94% CL, with one of these points continuing to fail at the 99.98% CL. An event that fails at the 99.98% CL is expected to occur only twice in 10 000 times, we observe it once in ten times, indicating that the null hypothesis is highly suspect.

To ensure that the data sample shown in Fig. 4 is not a statistical fluke, and to establish the repeatable and consistent absence of steady-state behavior, many more independent samples were generated. A total of 200 independent realizations were simulated, and the value of ξ_L^2 was measured for each realization at all of the times recorded in Fig. 4. The Gaussian distributed random variable $\overline{\xi_L^2}$ is constructed by splitting the 200 independent realizations into subsets of $N=20$ and calculating the average of each subset. Hence there are ten subsets, each one of the random variable $\overline{\xi_L^2}$ sampled at the ten times shown in Fig. 4, for a total of 100 realizations of $\overline{\xi_L^2}$ in the conjectured steady-state regime. We refer to each subset as a data set throughout the remainder of the manuscript. The $\chi^2_{\rm SS}$ test is applied independently to each of the ten data sets. The $T_{2(M-1)}$ test can be applied to any combination of two independent data sets. The T_{N-1} test is applied independently to each of the 100 realizations of $\overline{\xi_L^2}$. By conducting the statistical tests on the additional data sets we show that the results obtained for the original data set are systematic, as discussed below.

The $\chi^2_{\rm SS}$ statistic defined by Eq. (6) was calculated for each data set. The data sets all pass this test at greater than the 20% CL. Yet sufficiently many realizations of $\overline{\xi_L^2}$ fail the T_{N-1} test, to bring the null hypothesis of a steady-state into question. Fourteen of the 100 IID realizations of $\overline{\xi_L^2}$ fail the T_{N-1} test at the 90% CL, which in itself would not allow us to reach conclusions. However, as the CL criterion is tightened beyond acceptable standards a surprising number of points still fail. At the 99% CL, five of the 100 points fail. These five points also fail at the 99.8% CL. At the 99.98% CL, three points fail. Finally, at the 99.995% CL, one of the 100 points fails. The meaning of the 99.98% CL is that probability theory predicts the occurrence of two such events out of 10 000. Instead we observe three such events out of 100. Likewise the 99.995% CL corresponds to five events out of 100 000. We observe one such event out of 100. Extreme ''tail events'' thus occur with a frequency which is more than two orders of magnitude greater than the laws of probability would indicate.

We now turn to a discussion of the average asymptotic value of each data set of $M=10$ events (i.e., each set consists of the random variable $\overline{\xi_L^2}$ sampled at ten consecutive times in the asymptotic regime). The value of $\langle\overline{\xi_L^2}\rangle_M$ generated by each set of data is not consistent with the other sets. The ten values range from $\langle\overline{\xi_L^2}\rangle_M=33.05\pm1.31$ to 37.39 ± 1.34 [22]. Randomly picking pairs of data sets to compare using a $T_{2(M-1)}$ test, we find several instances where the null hy-

pothesis (that the two sets of data being compared were sampled independently from the same underlying distribution) fails at the 96% CL. Thus separate runs of a simulation frequently give statistically inconsistent results.

When all $N=200$ samples are combined into one data set, the average quantities are consistent with the empirical distribution function shown in Fig. 1: $\langle \overline{\xi^2_{L,N=200}} \rangle_M = 35.65 \pm 0.24$. In addition, this data set of the average over 200 IID samples passes the χ^2_{SS} test at the 97% CL. However, for the data sets of averages over $N=20$ IID realizations, the disparity in average asymptotic values obtained shows that sampling a subset of 20 IID samples is not consistent with the empirical distribution. The values for averages and variances obtained from the subsets are not unbiased estimators of the empirical distribution function. Furthermore, the statistical tests discussed so far reject the hypothesis of steady-state behavior, even at time scales orders of magnitude greater than the conjectured model relaxation time of $\tau \sim 10L^{z'}$. Thus there is no steady-state behavior for the data generated with random( ); instead there is an asymptotic dynamic behavior.

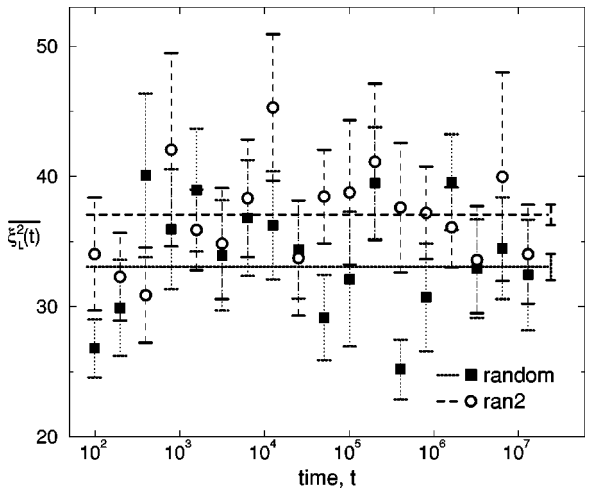

FIG. 5. The average width of the growth interface as obtained by 20 independent samples for each PRNG. The horizontal lines correspond to the respective average asymptotic values, $\langle \overline{\xi^2_L} \rangle_{M,\text{random}}$ and $\langle \overline{\xi^2_L} \rangle_{M,\text{ran2}}$, with the corresponding error bars plotted at the extreme end of each line.

B. Comparison of extrema with the steady-state distribution

In the preceding section the occurrence of many extreme tail events was established. In this present section we show that the occurrence of tail events can skew the average values obtained.

Comparing the data shown in Fig. 4 to the expected probability distribution, shown in Fig. 2, it is observed that the data points are skewed to the left side of the expected distribution. Eight of the ten points are below the mean value $\mu = 35.87$ and the lowest value, $\overline{\xi^2_{L\min}} = 25.18$, is in the left hand tail with less than 0.32% of the total area of the probability distribution function. The probability of obtaining a particular value for the minimum in a set of M IID random variables is described in Eq. (9). Using this formula we calculate the probability that the minimum of $M=10$ IID random variables drawn from the distribution shown in Fig. 2 is less than or equal to $\overline{\xi^2_{L\min}} = 25.18$ is only $p(x \leqslant x_{\min} = 25.18) = 3.6\%$.

Again we wish to determine if the first data set is a fluke event, so a statistical analysis based on all ten data sets obtained with random( ) is warranted. We find that three of the ten data sets have minima which come from the extreme left hand tail of the distribution, and $p(x \leqslant x_{\min}) < 6\%$ for each. As the probability distribution for the minimum, $p(x_{\min})$, is known (see Fig. 3), we can construct the probability for observing three such low probability events (from the left hand tail of the distribution) out of a total of ten events, and find this probability to be 1.7%. As such, we can state at the 98% CL that such results would not be obtained by random sampling.

A similar analysis can be carried out with respect to the maxima. In line with the observation that the data in Fig. 4 are skewed to the left side of the expected distribution, we find several values for maxima which are questionably low. The analogous probability for the maximum value in a set of M IID random variables is given by Eq. (10), and plotted in Fig. 3. Maximum values for four of the ten data sets are much smaller than expected, with each having probability $p(x \leqslant x_{\max}) < 12\%$. The probability of observing four such low probability events (from the left hand tail of the distribution) out of a total of ten events is 1.9%. Again we can state at the 98% CL that such results would not be obtained by random sampling.

C. Comparison of data from distinct PRNGs

To determine if the source of the observed asymptotic dynamic behavior resides in the PRNG random( ), results using a second PRNG, ran2( ), were also analyzed. ran2( ) is substantially slower than random( ), hence the comparison is based on 20 independent samples for each PRNG (i.e., we use only the initial set of data for random( ), shown in Fig. 4 [22]).

For the shortest length scale implemented ($L=127$), the time-series data for ran2( ) are self-consistent. The data pass the χ^2 test, the T_{N-1} test, and the tests for extremal values. However, at longer length scales the data for ran2( ) fail several statistical tests, making ran2( ) also suspect in simulations of BD. The tests performed at the longer length scales were adequate to show statistical inconsistencies [23], however, not with the high level of rigor demonstrated by the tests on the data at the shortest length scales [24].

A direct comparison of data generated by the two different PRNGs for $L=127$ is shown in Fig. 5. There are 18 sets of points that can be directly compared, including eight which were sampled at $t < \tau$. The $T_{2(M-1)}$ test for consistency between the two values at each time fails at the 90% level for three out of the 18 sets of points. Most striking is the direct comparison of average asymptotic values obtained for each PRNG, $\langle \overline{\xi^2_L} \rangle_{M,\text{random}} = 33.05 \pm 1.31$ and $\langle \overline{\xi^2_L} \rangle_{M,\text{ran2}} = 37.06 \pm 0.80$. A $T_{2(M-1)}$ test for the equivalence of the asymptotic averages of the data sets for random( ) and ran2( ) fails at the 99% confidence level (the exact probability of

failure is $p = 0.996$). Hence use of different PRNGs can yield statistically distinct values for averages.

The statistical tests reject the hypothesis that the two PRNGs sample the same underlying distribution, despite the fact that the asymptotic distributions shown in Fig. 1 agree. There is a unique steady-state distribution which is obtained in the limit of large numbers of independent samples. However, sampling of this distribution is nonstochastic, in that each sample average is not an unbiased estimator of the asymptotic distribution. Likewise the standard error for each sample does not lead to an unbiased estimator for the standard deviation of the asymptotic distribution.

V. DISCUSSION AND CONCLUSIONS

The original impetus for this study was an in-depth investigation of BD at long length and time scales. However, we encountered many features in the data that could not be easily explained; most notably, non-self-affine surface fluctuations. After searching for various corrections to scaling, which necessitated obtaining better statistics and exploration of longer times into the growth, the coupling to PRNGs became apparent, and motivated the detailed statistical analysis described in this manuscript.

It should be noted that there are discrepancies between values of the scaling exponents for BD reported in the literature [4,6,14,15]. At this point, we can only speculate that these discrepancies are due to the differences in the implementations of BD. Conclusions about the scaling exponents can only be reliably reached once difficulties with PRNs are resolved.

In retrospect, it is not surprising that the BD algorithm is more sensitive to correlations in PRN sequences than standard Monte Carlo (MC) simulations. In standard MC, comparison to the Boltzmann probability causes rejection of PRNs at pseudorandom points in the sequence; hence a second independent source of pseudorandomness influences the dynamics. In BD, all PRNs are used in the sequence produced (with the exception of the very few cases discussed in Sec. II B). Similarly, in restricted solid-on-solid models of growth (where physical constraints cause rejection of PRNs), the scaling exponents and the random walk distribution predicted by KPZ theory are recovered with great precision in numerical simulations [17,25].

We have demonstrated that computer implementations of BD can couple to certain PRNG algorithms. Results statistically inconsistent with general sampling assumptions and with the ergodic exploration of phase space were observed. Exploration of accessible phase space is not decoupled from the initialization of the PRNG. In addition, driving the dynamics of the system with different PRNGs results in sampling different areas of phase space. In conclusion, BD is a sensitive physical test of correlations in pseudorandom sequences. In general, PRNG algorithms can couple to models of stochastic, nonequilibrium phenomena. One must ensure that observed dynamical properties are inherent in the nonequilibrium model itself and not an artifact of coupling to PRNGs.

ACKNOWLEDGMENTS

The work at MIT was supported by the NSF Grant Nos. DMR-93-03667 and DMS-95-96217, and by DARPA Contract No. DABT63-95-C-0130. We have benefited from discussions with Y. Kantor, M. Smith, and N. Margolus.

[1] See, e.g., *Scale Invariance, Interfaces, and Non-equilibrium Dynamics*, edited by A. McKane, M. Droz, J. Vannimenus, and D. Wolf (Plenum, New York, 1995).

[2] M. Li and P. Vitányi, *An Introduction to Kolmogorov Complexity and Its Applications* (Springer-Verlag, New York, 1993); Y. Bar-Yam, *Dynamics of Complex Systems* (Addison-Wesley, Reading, MA, 1997).

[3] For recent reviews, see, e.g., *Dynamics of Fractal Surfaces*, edited by F. Family and T. Vicsek (World Scientific, Singapore, 1991); T. Halpin-Healy and Y.-C. Zhang, Phys. Rep. **254**, 215 (1995); A. L. Barabasi and H. E. Stanley, *Fractal Concepts in Surface Growth* (Cambridge University Press, Cambridge, England, 1995).

[4] J. Krug and H. Spohn, in *Solids Far From Equilibrium: Growth, Morphology and Defects*, edited by C. Godrèche (Cambridge University Press, Cambridge, England, 1991).

[5] M. J. Vold, J. Colloid Sci. **14**, 168 (1959).

[6] F. Family and T. Vicsek, J. Phys. A **18**, L75 (1985).

[7] M. Kardar, G. Parisi, and Y.-C. Zhang, Phys. Rev. Lett. **56**, 889 (1986).

[8] A. M. Ferrenberg, D. P. Landau, and K. Binder, J. Stat. Phys. **63**, 867 (1991).

[9] A. M. Ferrenberg and D. P. Landau, Phys. Rev. Lett. **69**, 3382 (1992).

[10] I. Vattulainen, T. Ala-Nissila, and K. Kankaala, Phys. Rev. Lett. **73**, 2513 (1994).

[11] L. N. Shchur, J. R. Heringa, and H. W. J. Blöte, Physica A **241**, 579 (1997).

[12] *Gnu C Library, Sun Release 4.1*. Original code is copyright Regents of the University of California, 1983.

[13] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*, 2nd ed. (Cambridge University Press, Cambridge, England, 1992).

[14] R. M. D'Souza, Int. J. Mod. Phys. C **8**, 941 (1997).

[15] P. Meakin, P. Ramanlal, L. M. Sander, and R. C. Ball, Phys. Rev. A **34**, 5091 (1986).

[16] B. Derrida and M. R. Evans, J. Phys. I **3**, 311 (1993).

[17] G. Foltin, K. Oerding, Z. Rácz, R. L. Workman, and R. K. P. Zia, Phys. Rev. E **50**, R639 (1994).

[18] Z. Rácz and M. Plischke, Phys. Rev. E **50**, 3530 (1994).

[19] W. Feller, *Introduction to Probability and Statistics, Vol. 1* (John Wiley & Sons, Inc., New York, 1957); R. S. Burington and D. C. May, Jr., *Handbook of Probability and Statistics with Tables* (McGraw-Hill Book Co., New York, 1970).

[20] ''Student,'' Biometrika **6**, 1 (1908); M. G. Bulmer, *Principles of Statistics* (Dover Publications, Inc., New York, 1979).

[21] J. Galambos, *The Asymptotic Theory of Extreme Order Statistics* (Robert E. Krieger Publishing Co., Malabar, FL, 1987).

[22] Note that the data set plotted in Fig. 4 is the set with the smallest asymptotic mean ($\langle\overline{\xi_L^2}\rangle_M = 33.05 \pm 1.31$). However, two additional data sets obtained with random( ) have asymptotic means below 34.0. We choose to plot the set shown in Fig. 4 as it was the first set obtained; as such we are consistent in comparing results with ran2( ), where only one data set was generated.

[23] R. M. D'Souza (unpublished).

[24] The theoretical fits discussed in Sec. III A are not precise enough to provide accurate measures of the cumulants of the empirical distribution function. Hence, the distribution for longer lengths cannot be assumed to be accurately described by a corresponding theoretical curve. Sampling sufficient data points to establish the empirical asymptotic distribution for the longer lengths is beyond the scope and intent of this manuscript.

[25] J. M. Kim and J. M. Kosterlitz, Phys. Rev. Lett. **62**, 2289 (1989).