Going through Rough Times: from Non-Equilibrium Surface Growth to Algorithmic Scalability

G. Korniss,[1] M.A. Novotny,[2] P.A. Rikvold,[3] H. Guclu,[1] and Z. Toroczkai[4]

[1]Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180-3590 USA

[2]Department of Physics and Astronomy and Engineering Research Center, Mississippi State University, Mississippi State, MS 39762-5167 USA

[3]School of Computational Science and Information Technology, Department of Physics, and Center for Materials Research and Technology, Florida State University, Tallahassee, FL 32306 USA

[4]Theoretical Division and Center for Nonlinear Studies, MS-B258 Los Alamos National Laboratory, Los Alamos, NM 87545 USA

ABSTRACT

Efficient and faithful parallel simulation of large asynchronous systems is a challenging computational problem. It requires using the concept of local simulated times and a synchronization scheme. We study the scalability of massively parallel algorithms for discrete-event simulations which employ conservative synchronization to enforce causality. We do this by looking at the simulated time horizon as a complex evolving system, and we identify its universal characteristics. We find that the time horizon for the conservative parallel discrete-event simulation scheme exhibits Kardar-Parisi-Zhang-like kinetic roughening. This implies that the algorithm is asymptotically scalable in the sense that the average progress rate of the simulation approaches a non-zero constant. It also implies, however, that there are diverging memory requirements associated with such schemes.

INTRODUCTION

Faithful and efficient simulation of complex systems with many interacting degrees of freedom is an important and challenging computational task. In a large class of systems the dynamic evolution is inherently stochastic and changes in the local configuration of the system occur randomly in space and time. The modeling of these systems can follow a "bottom-up" approach, starting with the definition of the "microscopic" dynamics. Then a master equation can be constructed with the corresponding transition probabilities. In most cases, for systems with many interacting degrees of freedom, even the numerical solution of the master equation (typically involving the numerical diagonalization of huge matrices) becomes insurmountable. This is when simulation becomes an invaluable tool for complex system modeling.

In the physics, chemistry, and biology communities these types of simulations are most commonly referred to as dynamic or kinetic Monte Carlo simulations. In computer science they are called discrete-event simulations. The updates (governed by the microscopic dynamics) in the (typically local) configuration of the system are considered discrete events. The basic notion of discrete-event simulation is that time is continuous and the discrete events occur instantaneously. Between events, the state (configuration) of the system remains unchanged. If the events occur at random instants of time, the dynamics can be

referred to as *asynchronous.* Examples of such systems include the Ising model with the Glauber or Metropolis dynamics (the discrete events are the spin-flip attempts), cellular communication networks (call arrivals), spatial epidemic models (infections), financial markets (buy/sell orders), or internet traffic (packet transmission/reception). In this paper we focus on systems with general but *short-ranged* interactions on *regular lattices* and assume that the event dynamics can be described as a superposition of a large number of independent Poisson processes running in parallel (*Poisson asynchrony*).

When the size of the system becomes large, parallelization may be needed to obtain results within an acceptable time frame. Massively parallel simulation for complex systems with asynchronous dynamics, i.e., *parallel discrete-event simulations* (PDES), is a standard technique among computer scientists. It is somewhat surprising that despite PDES having a long history as far as applications and scalability are concerned [1–2], very few of the PDES techniques have filtered through, e.g., to the physics community. Even the simplest random-site update Monte Carlo schemes [3] where update attempts converge to Poisson arrivals in the large system-size limit, were long believed to be inherently serial. In this regard, Lubachevsky's work [4–5] was rather illuminating by illustrating how to apply the PDES scheme to the Ising model with Glauber dynamics on a regular lattice [3, 6].

In a PDES scheme each processing element (PE) carries a subsystem of the full system via simple spatial decomposition. The difficulty of PDES is that the discrete events are not synchronized by a global clock since the dynamic is asynchronous. To put it simply, the paradoxical task is to (algorithmically) parallelize (physically) non-parallel dynamics. To achieve this, one must use the concept of local simulated times (or virtual times) and a synchronization scheme. The parallel algorithm must concurrently advance the local simulated time on each subsystem carried by a PE, without violating causality. In a "conservative" PDES scheme [7–8], only those PEs that are guaranteed not to violate causality are allowed to process their events and increment their local time. The rest of the PEs must "idle." In an "optimistic" approach [9], the PEs do not have to idle, but since causality is not guaranteed at every update, the simulated history on certain PEs can become corrupted. This requires a complex "rollback" protocol to correct erroneous computations. Both simulation approaches lead to an evolving and fluctuating time horizon during algorithmic execution.

For systems which can be modeled on regular lattices with short-ranged interactions, the conservative scheme can be highly efficient. Recently, it was implemented for modeling magnetization switching [10] and the dynamic phase transition [11] in large realizations of highly anisotropic thin-film ferromagnets. For example, the nearest-neighbor interaction implies that in order to ensure causality, PEs need to exchange their local simulated (or virtual) times only with "neighboring" PEs in the virtual PE topology. Communication times between PEs and possible idling due to the conservative synchronization protocol can be greatly suppressed by each PE carrying a large block of sites (spins) [4–5, 10], yielding encouraging efficiencies and utilizations (fraction of non-idling PEs). Since there is a finite number of PEs in any architecture, one can typically be satisfied with these "experimental" observations. However, as the number of PEs available to simulating complex systems increases to tens of thousands, scalability questions become fundamental. Further, questions, such as how the utilization behaves in the asymptotic limit when the number of PEs goes to infinity, truly lie at the heart of any PDES scheme.

We emphasize again, that the primary motivation for PDES is to perform parallel sim-

ulation for large systems *without* altering the original dynamics (easily implementable in serial). One may wonder why one should bother with the above complex parallel schemes when "trivial" or "embarrassing" parallelization would also be possible in principle. In that case one performs several (as many as the number of PEs) independent runs in parallel and aggregates the necessary statistics at the end. This approach is hard to beat in terms of programming effort and utilization (equal to one, since no synchronization, idling, and communication are needed). However, one often attempts to simulate *large* systems. In the worst case, the system to be simulated may *not* fit in the memory of a single CPU. In this case, clearly there is no alternative to PDES schemes. But even if a system fits in the CPU memory, unless supported by special harware, e.g., extended cache memory, the performance is gradually degraded as the cache is exceeded and the memory limit is approached [10]. In this case again, PDES is the method of choice for performance.

In this paper we address fundamental scalability questions for the general conservative parallel simulations for systems on regular lattices with short-range interactions. The way we tackle the problem, in some sense, goes opposite to the usual flow of a typical scientific modeling and simulation process. There, one develops advanced and sophisticated computational algorithms to study and understand systems in Nature. Here, by knowing how certain natural systems behave, we try to understand how advanced algorithms, PDES in particular, work: whether or not they are scalable, and how they can be optimized. Our approach is facilitated by a mapping [12–13] between non-equilibrium surface growth and the progress of the simulation (the evolution of the simulated time horizon). At the end, of course, based on the knowledge gained after answering fundamental scalability questions, one hopes to close this "loop" by devising and optimizing PDES schemes that can be used to investigate challenging problems in natural, artificial, or social systems. Analogies of a similar kind, e.g., that between phase transitions and computational complexity [14–15] also have turned out be highly fruitful to gain more insight in traditionally difficult problems in computer science. Also, exploiting analogies between the evolution of the time horizon and that of known *physical* systems appears to be rather helpful in understanding the performance for optimistic schemes as well. There is some evidence [16–17] that the time horizon in rollback-based schemes can exhibit self-organized criticality and power-law spatio-temporal correlations, which can be crucial to extract the scalability properties.

Based on our mapping between non-equilibrium surface growth and the progress of the simulation, it is clear that the tools and frameworks of modern statistical physics, in particular those of non-equilibrium interface/surface growth [18–20], can be extremely helpful in analyzing and understanding the asymptotic scalability properties of PDES schemes. To this end, one must look at the simulation scheme itself as an evolving interacting system of individual PEs where the synchronization rules among the PEs constitute the effective interaction. The evolution of this simulated time horizon, in particular its average progression rate and statistical spread, will determine the scalability properties of the corresponding PDES scheme.

In the next section we give an overview of the basic conservative PDES scheme [4–5]. Then we characterize the morphological properties of the evolving random surface associated with the simulated time horizon. These findings yield direct implications for the scalability of the conservative algorithm for PDES.

THE BASIC CONSERVATIVE APPROACH

We consider a d-dimensional hypercubic regular lattice topology, where the underlying physical system has only nearest-neighbor interactions. However, our results hold for any short-range regular interaction pattern. In this paper we consider the case of simple Poisson asynchrony. Update attempts at each site are identical and independent Poisson processes (thus, the random simulated time increments between two successive update attempts are exponentially distributed) and are also *independent of the state of the underlying physical system.* The consequence of the former is that the evolution of the simulated time horizon completely *decouples* from the behavior and evolution of the underlying physical system. For simplicity, we discuss in detail the "worst-case" scenario, in which each PE carries one site (e.g., one spin). In this basic conservative scheme, each PE generates its own local simulated time for the next update attempt. (The actual update probabilities depend on the underlying systems, e.g., through energetics, but they do not affect the evolution of the time horizon.)

The set of local simulated times for N PEs, $\{\tau_i(t)\}_{i=1}^{N}$, constitutes the simulated time horizon. Here t is the discrete number of parallel steps simultaneously performed on each PE directly related to real/wall-clock time, or if the architecture operates in an asynchronous execution mode, t is simply the continuous real time. On a regular d-dimensional hypercubic lattice $N = L^d$, where L is the linear size of the lattice. In physics applications one typically specifies the initial configuration (i.e., at $\tau = 0$) of the underlying physical system. This translates to $\tau_i(0) = 0$ for every site for the initial condition of the parallel simulation. Then at each parallel update, *only* those PEs for which the local simulated time is *not greater* than the local simulated times of their nearest neighbors, can increment their local time by an exponentially distributed random amount, $\eta_i(t)$. The other PEs must idle. Without loss of generality we take independent, identically distributed (iid) exponential variables of mean one in simulated time units (stu), $\langle \eta_i(t) \rangle = 1$. Due to the continuous nature of the random simulated times, for $t > 0$ the probability of equal-time updates for any two sites is of measure zero. The comparison with nearest-neighbor simulated times and, if necessary, idling enforces causality. Also, at worst, the PE with the global minimum simulated time can make progress, so the algorithm is free from deadlock. For this basic conservative scheme, the theoretical efficiency or utilization (ignoring communication overheads) is simply the (average) fraction of non-idling PEs. This corresponds to the density of local minima of the simulated stochastic time horizon, which determines the average progress rate of the simulation. It can be written as

$$\langle u(t) \rangle_N = \left\langle \sum_{i=1}^{N} \prod_{j \in D_i^{\mathrm{nn}}} \Theta\left(\tau_j(t) - \tau_i(t)\right) \right\rangle , \qquad (1)$$

where D_i^{nn} is the set of nearest neighbors of site i, $\Theta(\cdot)$ is the Heaviside step function, and the $\langle \ldots \rangle$ denotes an ensemble average, i.e., an average over many independent simulations.

Another important aspect of the simulation is the width of the distribution of the local simulated times. This property can have serious effects on the "measurement part" of the algorithm, e.g. when one attempts to collect and compute simple statistics for the full underlying physical system "on the fly." Therefore, one must determine the *statistical spread* (width) of the time horizon as was pointed out in Ref. [21]. This quantity can be
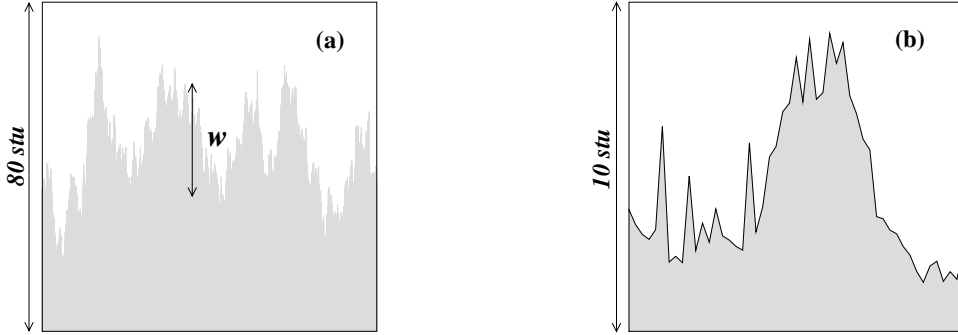
Figure 1: (a) Snapshot configuration of the actual simulated time horizon for the one-dimensional one site per PE topology with L=10,000 PEs. The time horizon propagates "upwards" in the figure, and w indicates the average width of the time horizon. (b) A short segment (50 sites) of the same configuration. The arrows on the left indicate the vertical scales in simulated time units (stu).

characterized by

$$\left\langle w^2(t) \right\rangle_N = \left\langle \frac{1}{N} \sum_{i=1}^{N} [\tau_i(t) - \bar{\tau}(t)]^2 \right\rangle , \qquad (2)$$

where $\bar{\tau}(t)$=$(1/N)\sum_{i=1}^{N} \tau_i(t)$ is the mean progress ("height") of the time horizon. The behavior of the width, $\langle w^2(t) \rangle$, alone typically captures and identifies the universality class of the non-equilibrium growth process [18–20].

EVOLUTION OF THE SIMULATED TIME HORIZON

The conservative synchronization protocol together with the virtual communication topology of the PEs (mimicking the interaction topology of the underlying physical system) fully specify the "microscopic" dynamics of the growth process associated with the evolution of the time horizon. The rules of the conservative synchronization can be compactly summarized as

$$\tau_i(t+1) = \tau_i(t) + \prod_{j \in D_i^{\mathrm{nn}}} \Theta \left(\tau_j(t) - \tau_i(t) \right) \eta_i(t) . \qquad (3)$$

The above equation simply reflects that the local simulated time at site (PE) i is incremented by an exponentially distributed random amount $\eta_i(t)$, *provided* that site i is a local minimum of the time horizon. This stochastic time evolution equation is exact for the basic conservative scheme and can be easily simulated on a serial (!) workstation. Snapshots of the evolving and fluctuating time horizon obtained from direct simulations are shown in Fig. 1. We discuss in detail the one-dimensional case (N=L) with periodic boundary conditions (i.e., ring topology). Replacing the Heaviside step function with a limiting smooth representation, one can perform standard coarse graining on Eq. (3) [12], yielding

$$\frac{\partial \hat{\tau}}{\partial t} = \frac{\partial^2 \hat{\tau}}{\partial x^2} - \lambda \left(\frac{\partial \hat{\tau}}{\partial x} \right)^2 + \hat{\eta}(x,t) . \qquad (4)$$

Here $\hat{\tau}$ is the coarse-grained "height" fluctuation, and the temporal and spatial derivatives are just the naïve continuum interpretations of the differences, e.g., $\partial \hat{\tau}/\partial x$=$\tau_i - \tau_{i-1}$. Similarly, $\hat{\eta}$ is the coarse-grained noise, delta correlated in space and time. Higher order terms

in Eq. (4) are dropped since they are irrelevant in a Renormalization Group (RG) sense. Equation (4) is the Kardar-Parisi-Zhang (KPZ) equation [22], which has turned out to be of central importance and have a wealth of applications for numerous artificial and natural growth processes over the past two decades, including molecular beam epitaxy, electrochemical deposition, fluid flow in porous media, and growth of bacterial colonies [18–19]. Thus, we expect that the simulated time horizon exhibits *kinetic roughening*, the main feature of KPZ growth. Indeed, our direct simulation for the time horizon evolution, Eq. (3), confirms the coarse-graining approach [Fig. 2(a) and (b)]. There is a system-size dependent characteristic time scale, the crossover time, $t_\times \sim L^z$. For very early times, mostly microscopic details of the dynamics influence the width. For intermediate times, while $t \ll t_\times$ still, the width grows as a power law $\langle w^2(t) \rangle_L \sim t^{2\beta}$, where β is the *growth exponent*. For late times, $t \gg t_\times$, the width *saturates* for any finite system size. In this regime the surface reaches a steady-state evolution, and the fluctuations about the mean are stationary. The saturation or steady-state value of the width, however, scales as a power law with the system size, $\langle w^2(\infty) \rangle_L \sim L^{2\alpha}$, where α is the *roughness exponent*. The time horizon through its progress exhibits exactly the above scaling behavior with $\beta \simeq 1/3$ and $\alpha \simeq 1/2$, consistent with the exact one-dimensional KPZ exponents [18–19, 22] [Fig. 2(a)]. This type of temporal and system-size scaling is consistent with the dynamic scaling hypothesis [18, 23] and can be expressed through the Family-Vicsek scaling relation

$$\left\langle w^2(t) \right\rangle_L = L^{2\alpha} f(t/L^z) \tag{5}$$

together with the important scaling law, $\alpha = \beta z$. Note that the scaling function $f(x)$ depends on t and the linear system size L only through the specific combination t/L^z, reflecting the importance of the crossover time $t_\times$. For small values of its argument $f(x)$ behaves as a power law, while for large arguments it approaches a constant

$$f(x) \sim \begin{cases} x^{2\beta} & \text{if } x \ll 1 \\ \text{const.} & \text{if } x \gg 1 \end{cases} . \tag{6}$$

The existence of the above scaling function implies that if one plots the rescaled variables $\langle w^2(t) \rangle_L / L^{2\alpha}$ vs t/L^z, then curves for different system sizes collapse for intermediate and late times. We have confirmed this data collapse for the simulated time horizon [Fig. 2(b)].

Direct simulation results for the average rate of progress of the time horizon (equivalent to the utilization in the PDES algorithm) for various system sizes are shown in Fig. 2(c). The utilization $\langle u(t) \rangle_L$ decreases monotonically with time towards a long-time asymptotic limit well separated from zero, $\langle u(\infty) \rangle_\infty \approx 0.25$. The fact that it cannot vanish in the infinite system-size limit can be argued based on an important universal feature of KPZ-like surfaces: The steady-state KPZ surface in one dimension is governed by the Edwards-Wilkinson Hamiltonian [24], i.e., it is essentially a random-walk profile [18]. At coarse-grained length scales the local *slopes* become *independent*, yielding a *non-zero* average density of local minima, i.e., a non-zero average rate of progress of the simulation in the $L \to \infty$ limit in the steady-state.

In higher dimensions we observe the same qualitative behavior as for $d=1$ [13]. The surface roughens and *saturates* for any finite system, as seen in Fig. 3(a) and (c). Simultaneously, the density of local minima decreases monotonically towards its asymptotic ($t \to \infty$) finite-size value [Fig. 2(b) and (d)]. Again, the steady-state density of local minima appears
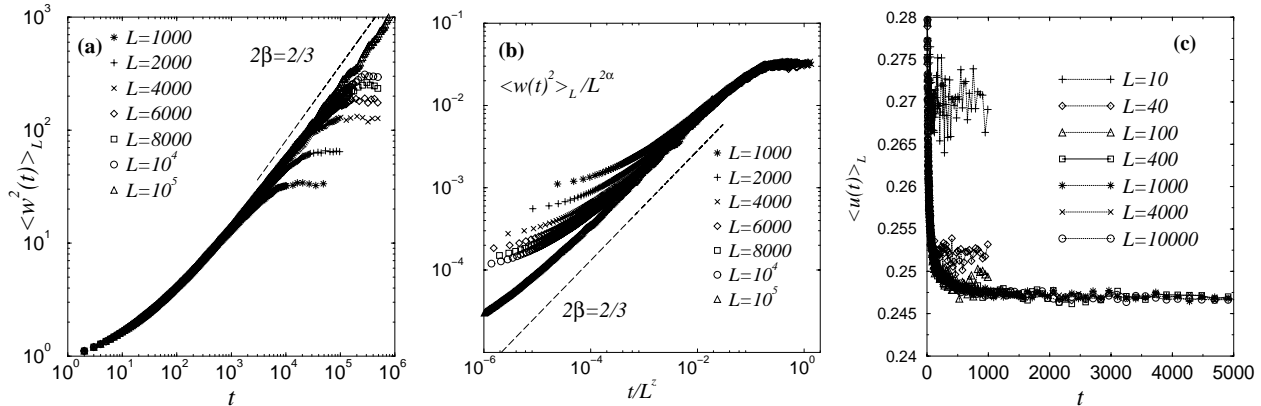
Figure 2: (a) Kinetic roughening of the simulated time horizon for the one-dimensional one site per PE basic conservative scheme. Note the log-log scales, indicating the power-law growth of the width before saturation. The dashed line corresponds to a power law with the exact KPZ exponent $2\beta=2/3$. (b) The same behavior as in (a), using rescaled variables to demonstrate the dynamic scaling hypothesis, Eq. (5). (c) Time dependent utilization for various system sizes.

to be well separated from zero. For $d=2$ $\langle u(\infty)\rangle_\infty \approx 0.12$, and for $d=3$ $\langle u(\infty)\rangle_\infty \approx 0.075$. The $\langle u(\infty)\rangle_\infty \sim \mathcal{O}(1/K)$ behavior appears to be rather general [21], where $K=2d$ is the number of nearest neighbors on a regular lattice. Similar to the $d=1$ case, corrections to scaling are very strong, both for the surface width and the density of local minima. While for $d=1$ we were able to simulate large systems ($L \gg 10^3$) to obtain the KPZ scaling exponents and the steady-state finite-size behavior of $\langle u(\infty)\rangle_L$, in higher dimensions the relatively small system sizes prevented us from extracting the scaling behavior of the width and the finite-size effects of the density of local minima. We conjecture that the simulated time horizon exhibits KPZ-like evolution in higher dimensions as well.

Next, we investigate in detail the steady-state scaling properties of the simulated time horizon for $d=1$, which directly translates to the asymptotic scalability properties of the corresponding PDES.

SCALING AND SCALABILITY

First, we provide further numerical evidence that the time horizon belongs to the KPZ universality class, in particular, in one dimension in the steady state, it is governed by the Edwards-Wilkinson (EW) Hamiltonian [24],

$$\mathcal{H}_{\text{EW}} \propto \frac{1}{2}\int dx (\partial\hat{\tau}/\partial x)^2. \tag{7}$$

Then for small wave-vectors k, the average steady-state structure factor of the surface should behave as

$$S(k) = \langle \tau_k \tau_{-k}\rangle/L \propto 1/k^2 , \tag{8}$$

where $\tau_k = \sum_{j=1}^{L} e^{-ikj}\tau_j$ is the spatial Fourier transform of the time horizon. This expectation is confirmed by direct simulations as shown in Fig. 4(a). The above small-k behavior of the structure factor also governs the steady-state scaling of the width, $\langle w^2(\infty)\rangle_L \sim L$, implying $\alpha=1/2$ in one dimension [Fig. 4(b)]. In addition to measuring averages, we also constructed
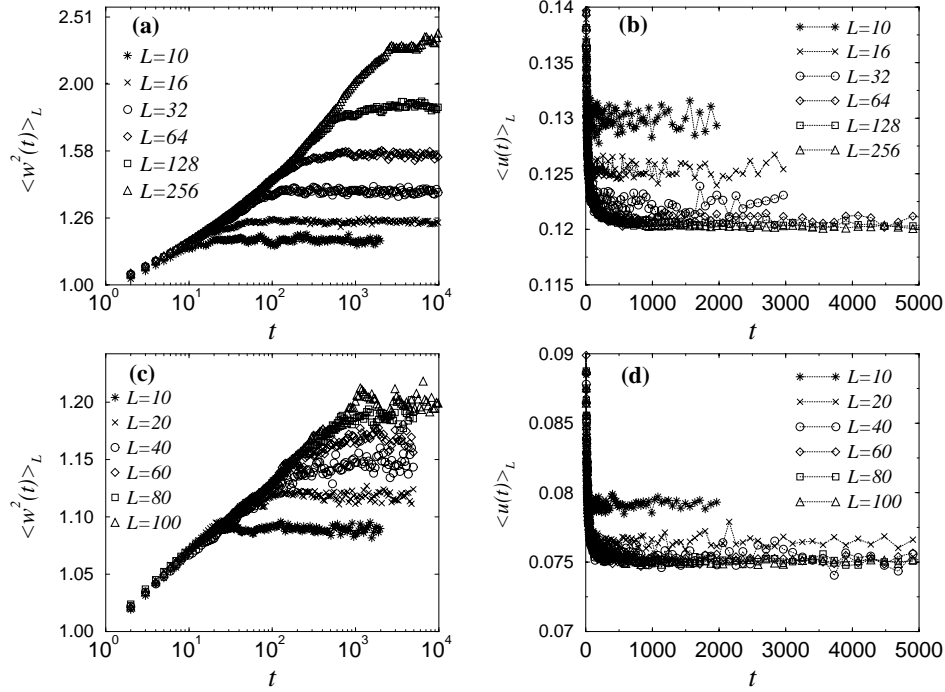
Figure 3: Evolution of the simulated time horizon in $d=2$ and $d=3$: (a) surface width in $d=2$; (b) density of local minima in $d=2$; (c) surface width in $d=3$; (d) density of local minima in $d=3$.

the *full* width distribution in the steady state, a universal signature of rough surfaces. By comparing it to existing analytic results [25], we confirmed [12] that the time horizon surface in one dimension belongs to the EW class.

Based on our results that the simulated time horizon exhibits KPZ-like kinetic roughening, we now address the implications for the scalability of conservative PDES schemes. We already argued that KPZ universal surfaces evolve toward a steady state where the coarse-grained local slopes, $\partial\hat{\tau}/\partial x$, become independent [see Eq. (7)]. This ensures that there is a finite density of local minima, so that the PDES algorithm progresses at a non-vanishing rate in the limit of an infinite number of PEs. Just as important for the utilization are the finite-size effects. Using exact calculations based on the EW Hamiltonian in one dimension [26] and scaling arguments in higher dimensions [27], one can obtain the universal finite-size effects for the growth rate of generic KPZ-like processes

$$\langle u(\infty)\rangle_L \simeq \langle u(\infty)\rangle_\infty + \frac{\text{const.}}{L^{2(1-\alpha)}} \ . \tag{9}$$

The above equation can be used to estimate the utilization (average rate of progress of the simulation) and to extrapolate to the value for the infinite number of PEs, $\langle u(\infty)\rangle_\infty$. Equation (9) for the finite-size effects is in full agreement with the simulations [Fig. 5(a)] and yields $\langle u(\infty)\rangle_\infty$=0.246461(7) for the one dimensional case (α=1/2). While the actual asymptotic value of the density of local minima depends on "microscopic" measures, whether its asymptotic value vanishes or not, is fully governed by macroscopic characteristics and the corresponding universality class [12, 26]. We emphasize again that for the KPZ class this asymptotic value is *non-zero*. The above "worst-case scenario" small values for the utilization, e.g., $\langle u(\infty)\rangle_\infty$≈0.12 for the two-dimensional one site per PE case, should not
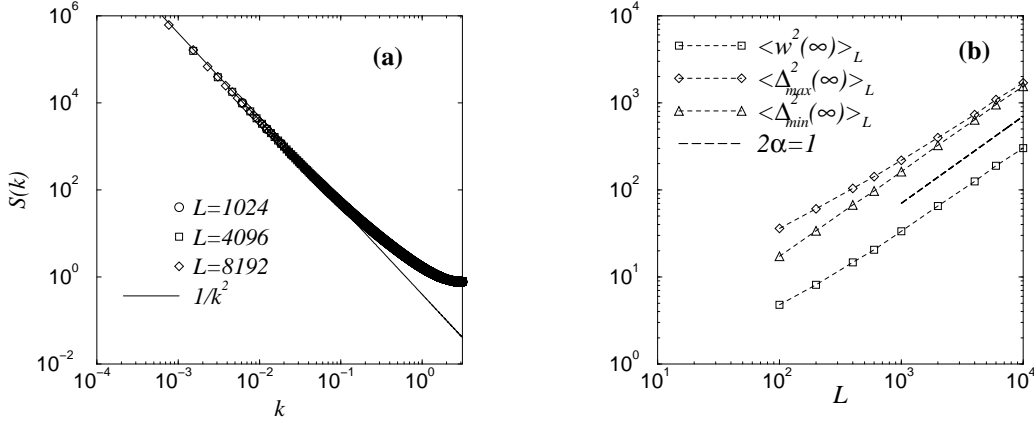
Figure 4: (a) Steady-state structure factor of the simulated time horizon. The straight solid line corresponds to the theoretical prediction Eq. (8). (b) Steady-state average width and extremal fluctuations of the time horizon. The dashed line corresponds to a power law with the exact KPZ/EW exponent $2\alpha=1$. Both graphs are for the one-dimensional one site per PE basic conservative scheme.

be too discouraging: in an actual implementation for large systems, one typically chooses a large (linear) block size l, carried by each PE. Then the favorable small value of the "surface-to-volume" ratio for the blocks (which scales as $1/l$) significantly increases the efficiency or utilization (fraction of non-idling PEs) by reducing the density of events for which synchronization (and possible idling) is needed [Fig. 5(b)].

The above findings for the utilization or density of local minima, which determines the average rate of progress of the simulation, imply that the "simulation part" of the conservative scheme is scalable. That is, if we run the simulation for long times, the average progress rate approaches a constant. However, the kinetic roughening exhibited by the time horizon has a disturbing implication: in the steady state the width (spread) of the simulated time horizon *diverges* with the number of PEs as

$$\langle w^2(\infty)\rangle_L \sim L^{2\alpha} \; . \tag{10}$$

This scaling behavior for large L is also confirmed by simulations [Fig. 4(b)], and it is contrary to the conclusions of Ref. [21]. This property creates an additional difficulty for collecting statistics (e.g., to perform simple averages) "on the fly" during the course of the simulation. The diverging width means that the memory requirement *per PE*, for temporarily storing (buffering) data, diverges as we increase the number of PEs. In this sense we may call the "measurement part" of the bare conservative scheme asymptotically nonscalable. Thus, in an actual application, the programmer must implement some global synchronization or a moving "window" with respect to the global minimum of the time horizon [28]. However such a "window" can have negative effects on the large-L utilization.

Along these lines of questioning, we are also interested in the *extremal* fluctuations of the time horizon. Namely, what is the typical size of the largest fluctuations above and below the mean, $\Delta_{\max}(t)\equiv(\tau_{\max}(t) - \bar{\tau}(t))$ and $\Delta_{\min}(t)\equiv(\bar{\tau}(t) - \tau_{\min}(t))$, where $\tau_{\max}(t)$ and $\tau_{\min}(t)$ are the *global* maximum and minimum simulated times among L PEs, respectively. We found that in the steady state

$$\left\langle \Delta_{\max}^2(\infty)\right\rangle_L \sim \left\langle \Delta_{\min}^2(\infty)\right\rangle_L \sim L^{2\alpha} \; , \tag{11}$$
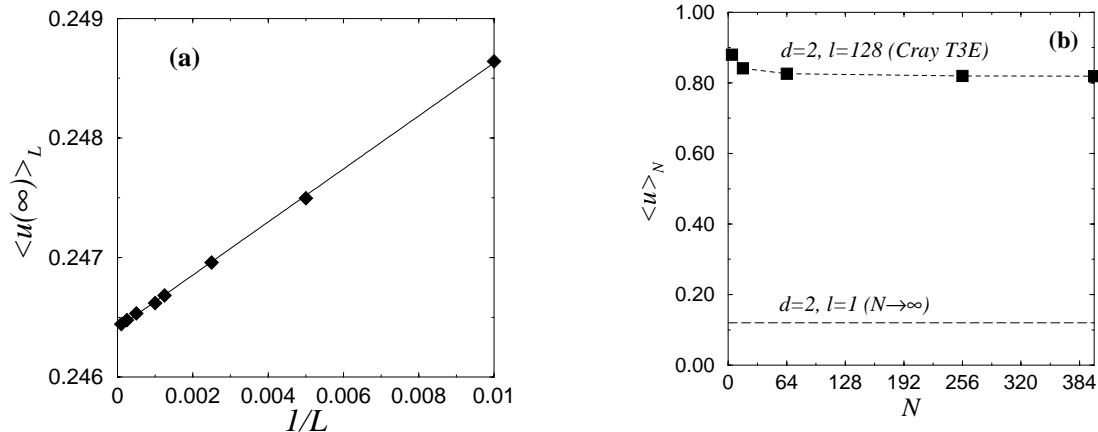
Figure 5: (a) Steady-state utilization (average rate of progress) for the one-dimensional one site per PE basic conservative scheme as function of the inverse system size. (b) Utilization (solid squares) for the actual implementation of the two-dimensional square-lattice kinetic Ising simulation on a Cray T3E up to N=400 PEs [10]. Each PE carries a 128×128 block of spins. The dashed horizontal line indicates the asymptotic limit for the two-dimensional one site per PE case.

i.e., they scale the *same way* as the width $\langle w^2(\infty) \rangle_L$ [Fig. 4(b)]. This should not come as a surprise, since this surface is highly correlated, dominated by long-wavelength fluctuations spreading over macroscopic length scales. This finding, again, is consistent with the extremal fluctuations found for general KPZ surfaces [29].

The above universal characteristics hold for the sensible and more efficient many sites per PE case and/or when the interaction and the corresponding communication pattern extends beyond nearest-neighbor PEs (but is still short ranged with a finite cutoff). For the case of many sites per PE, however, the saturation occurs at a later time and the time horizon exhibits a substantially larger width [30].

SUMMARY AND OUTLOOK

We have studied the statistical properties of the basic conservative parallel scheme for regular lattice topologies. We found that the evolution of the simulated time horizon belongs to the well-known KPZ dynamic universality class of non-equilibrium surfaces. This type of growth is characterized by a *non-zero* density of local minima, implying a non-zero rate of propagation in the limit of an infinite number of PEs. We also determined the asymptotic finite-size corrections to this constant when the number of PEs is large but finite. Thus, the "simulation" part of the algorithm is scalable. Further, we showed that the spread (width) of the time horizon approaches a finite constant for a *finite* number of PEs, but this constant *diverges* as a power law in the limit of an infinite number of PEs. The same holds for the extremal fluctuations above and below the mean. This "macroscopic" roughness of the simulated time horizon means that the "measurement part" of the bare conservative scheme is not scalable. That is, there is an extra difficulty associated with taking statistical measurements "on the fly." Intermittent data on each PE have to be stored until all PEs reach the simulated time instant at which some statistics collection (e.g., simple averaging over the full physical application) is to be performed. The diverging spread of the time horizon implies a diverging storage need for this purpose on every PE. Thus, the programmer

has to implement some global synchronization or windowing technique to limit the spread of the simulated time horizon in order not to exceed the memory constraint. By knowing exactly the finite-size dependence of the spread, for fixed L one can determine the optimal time between global synchronizations or the optimal window size.

Our findings are universal in the sense that they hold for any *short-range* "interaction" topology for PEs on regular lattices. They are also valid in the case when each PE carries a block of sites. The *asymptotic* scaling behavior is again governed by the KPZ exponents, in such a way that for larger and larger blocks, there is a crossover from the almost "random deposition" [18] to KPZ-like growth at a later and later time.

We must mention that there have been earlier attempts to theoretically describe the scalability of the basic conservative PDES scheme. To obtain an analytically tractable scalability model, Greenberg et al. [21] introduced the K-random model. In this model at each update attempt, PEs compare their local simulated times to the local simulated times of K *randomly* chosen PEs (rechosen at every update attempt). They showed that in the $t \to \infty$, $N \to \infty$ limit the average rate of progress of the simulation converges to a *non-zero* constant, $1/(K+1)$. They also showed that the evolution of the time horizon converges to a traveling-wave solution described by a *finite width* of the distribution of the local times. Finally, they suggested that convergence to a traveling-wave solution in the $t \to \infty$, $N \to \infty$ limit is universal and applicable for regular lattices as well. In obtaining this conclusion they made the assumption that replacing the "interaction" between nearest-neighbor PEs on a regular grid with the same interaction between *randomly* chosen PEs does not change the universality class of the time horizon. However, it does. Comparing the local time for each PE to K randomly chosen others essentially turns the model in to a mean-field-like one where the time surface is short-range correlated and has a finite width in the limit of an infinite number of PEs. As we have shown in this paper, the time horizon of the conservative PDES for regular lattices and short-ranged interactions with finite cutoffs exhibits KPZ-like kinetic roughening. This shows that the underlying communication topology of the PEs has crucial effects on the "universal" characteristics of the simulated time horizon.

Realizing the importance of the communication topology of the PEs, we are currently investigating how to turn the original conservative scheme on *regular lattices* into a fully scalable one, where both the "simulation" and the "measurement" parts are scalable, without the need for any global synchronization or windowing technique. Results of these studies will be published elsewhere.

ACKNOWLEDGMENTS

REFERENCES

1. R. Fujimoto, *Commun. of the ACM* **33**, 30 (1990).
2. D.M. Nicol and R.M. Fujimoto, *Annals of Operations Research* **53** 249 (1994).
3. K. Binder and D.W. Heermann, *Monte Carlo Simulation in Statistical Physics. An Introduction*, 3rd ed. (Springer, Berlin, 1997).

4. B.D. Lubachevsky, *Complex Systems* **1**, 1099 (1987).

5. B.D. Lubachevsky, *J. Comput. Phys.* **75**, 103 (1988).

6. R.J. Glauber, *J. Math. Phys.* **4**, 294 (1963).

7. K.M. Chandy and J. Misra, *IEEE Trans. on Softw. Eng.* **SE-5**, 440 (1979).

8. K.M. Chandy and J. Misra, *Commun. ACM* **24**, 198 (1981).

9. D.R. Jefferson, *Assoc. Comput. Mach. Trans. Programming Languages and Systems* **7**, 404 (1985).

10. G. Korniss, M.A. Novotny, and P.A. Rikvold, *J. Comput. Phys.* **153**, 488 (1999).

11. G. Korniss, C.J. White, P.A. Rikvold, and M.A. Novotny, *Phys. Rev. E* **63**, 016120 (2001).

12. G. Korniss, Z. Toroczkai, M.A. Novotny, and P.A. Rikvold, *Phys. Rev. Lett.* **84**, 1351 (2000).

13. G. Korniss, M.A. Novotny, Z. Toroczkai, and P.A. Rikvold, in *Computer Simulation Studies in Condensed Matter Physics XIII*, Springer Proceedings in Physics, Vol. 86, editors D.P. Landau, S.P. Lewis, and H.-B. Schüttler (Springer-Verlag, Berlin, Heidelberg, 2001), pp. 183–188.

14. *Frontiers in Problem Solving: Phase Transitions and Complexity*, edited by T. Hogg, B.A. Huberman, and C. Williams, Artif. Intell. **81** issue 1-2 (1996).

15. R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, *Nature (London)* **400**, 133 (1999).

16. B.J. Overeinder, *Distributed Event-driven Simulation: Scheduling Stategies and Resource Management*, Ph.D. thesis, Universiteit van Amsterdam (2000).

17. B.J. Overeinder, A. Schoneveld, and P.M.A. Sloot, *Proceedings of the 15th Workshop on Parallel and Distributed Simulation* (IEEE Comput. Soc., Los Alamitos, CA, 2001) pp. 145–152.

18. A.-L. Barabási and H.E. Stanley, *Fractal Concepts in Surface Growth* (Cambridge University Press, Cambridge, 1995).

19. Timothy Halpin-Healy and Yi-Cheng Zhang, *Phys. Rep.* **254** 215 (1995).

20. J. Krug, *Adv. Phys.* **46**, 139 (1997).

21. A.G. Greenberg, S. Shenker, and A.L. Stolyar, *Proc. ACM SIGMETRICS, Performance Eval. Rev.* **24**, 91 (1996).

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

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

24. S.F. Edwards and D.R. Wilkinson, *Proc. R. Soc. London, Ser A* **381**, 17 (1982).

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

26. Z. Toroczkai, G. Korniss, S. Das Sarma, and R.K.P. Zia, *Phys. Rev. E* **62**, 276 (2000).

27. J. Krug and P. Meakin, *J. Phys. A* **23**, L987 (1990).

28. G. Korniss, M.A. Novotny, A.K. Kolakowska, and H. Guclu, "Statistical Properties of the Simulated Time Horizon in Conservative Parallel Discrete-Event Simulations", accepted for the *Proceedings of ACM Symposium On Applied Computing*, Madrid, Spain (2002).

29. S. Raychaudhuri, M. Cranston, C. Przybyla, and Y. Shapir, *Phys. Rev. Lett.* **87**, 136101 (2001).

30. A.K. Kolakowska, M.A. Novotny, and G. Korniss, unpublished.

References

[1] R. Fujimoto, *Commun. of the ACM* **33**, 30 (1990).

[2] D.M. Nicol and R.M. Fujimoto, *Annals of Operations Research* **53** 249 (1994).

[3] K. Binder and D.W. Heermann, *Monte Carlo Simulation in Statistical Physics. an Introduction*, 3rd ed. (Springer, Berlin, 1997).

[4] B.D. Lubachevsky, *Complex Systems* **1**, 1099 (1987).

[5] B.D. Lubachevsky, *J. Comput. Phys.* **75**, 103 (1988).

[6] R.J. Glauber, *J. Math. Phys.* **4**, 294 (1963).

[7] K.M. Chandy and J. Misra, *IEEE Trans. on Softw. Eng.* **SE-5**, 440 (1979).

[8] K.M. Chandy and J. Misra, *Commun. ACM* **24**, 198 (1981).

[9] D.R. Jefferson, *Assoc. Comput. Mach. Trans. Programming Languages and Systems* **7**, 404 (1985).

[10] G. Korniss, M.A. Novotny, and P.A. Rikvold, *J. Comput. Phys.* **153**, 488 (1999).

[11] G. Korniss, C.J. White, P.A. Rikvold, and M.A. Novotny, *Phys. Rev. E* **63**, 016120 (2001).

[12] G. Korniss, Z. Toroczkai, M.A. Novotny, and P.A. Rikvold, *Phys. Rev. Lett.* **84**, 1351 (2000).

[13] G. Korniss, M.A. Novotny, Z. Toroczkai, and P.A. Rikvold, in *Computer Simulation Studies in Condensed Matter Physics XIII*, Springer Proceedings in Physics, Vol. 86, editors D.P. Landau, S.P. Lewis, and H.-B. Schüttler (Springer-Verlag, Berlin, Heidelberg, 2001), p. 183–188.

[14] *Frontiers in Problem Solving: Phase Transitions and Complexity*, edited by T. Hogg, B.A. Huberman, and C. Williams, Artif. Intell. **81** issue 1-2 (1996).

[15] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, *Nature (London)* **400**, 133 (1999).

[16] B.J. Overeinder, *Distributed Event-driven Simulation: Scheduling Stategies and Resource Management*, Ph.D. thesis, Universiteit van Amsterdam, (2000).

[17] B.J. Overeinder, A. Schoneveld, and P.M.A. Sloot, *Proceedings of the 15th Workshop on Parallel and Distributed Simulation* (IEEE Comput. Soc., Los Alamitos, CA, 2001) p. 145–152.

[18] A.-L. Barabási and H.E. Stanley, *Fractal Concepts in Surface Growth* (Cambridge University Press, Cambridge, 1995)

[19] Timothy Halpin-Healy and Yi-Cheng Zhang, *Phys. Rep.* **254** 215 (1995).

[20] J. Krug, *Adv. Phys.* **46**, 139 (1997).

[21] A.G. Greenberg, S. Shenker, and A.L. Stolyar, *Proc. ACM SIGMETRICS, Performance Eval. Rev.* **24**, 91 (1996).

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

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

[24] S. F. Edwards and D. R. Wilkinson, *Proc. R. Soc. London, Ser A* **381**, 17 (1982).

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

[26] Z. Toroczkai, G. Korniss, S. Das Sarma, and R.K.P. Zia, *Phys. Rev. E* **62**, 276 (2000).

[27] J. Krug and P. Meakin, *J. Phys. A* **23**, L987 (1990).

[28] G. Korniss, M.A. Novotny, A.K. Kolakowska, and H. Guclu, "Statistical Properties of the Simulated Time Horizon in Conservative Parallel Discrete-Event Simulations", accepted for the *Proceedings of ACM Symposium On Applied Computing*, Madrid, Spain (2002).

[29] S. Raychaudhuri, M. Cranston, C. Przybyla, and Y. Shapir, *Phys. Rev. Lett.* **87**, 136101 (2001).

[30] A.K. Kolakowska, M.A. Novotny, and G. Korniss, unpublished.