ITR: Non-equilibrium surface growth and the scalability of parallel discrete-event simulations for large asynchronous systems

NSF DMR-0113049

http://www.rpi.edu/~korniss/Research/gk_research.html

PIs: Gyorgy Korniss (Rensselaer), Mark Novotny (Mississippi State U.)

postdoc: Alice Kolakowska (Mississippi State U.)

graduate student: H. Guclu (Rensselaer)

undergraduate students: Katie Barbieri, John Marsh, Brad McAdams (Rensselaer); Shannon Wheeler (MSState)

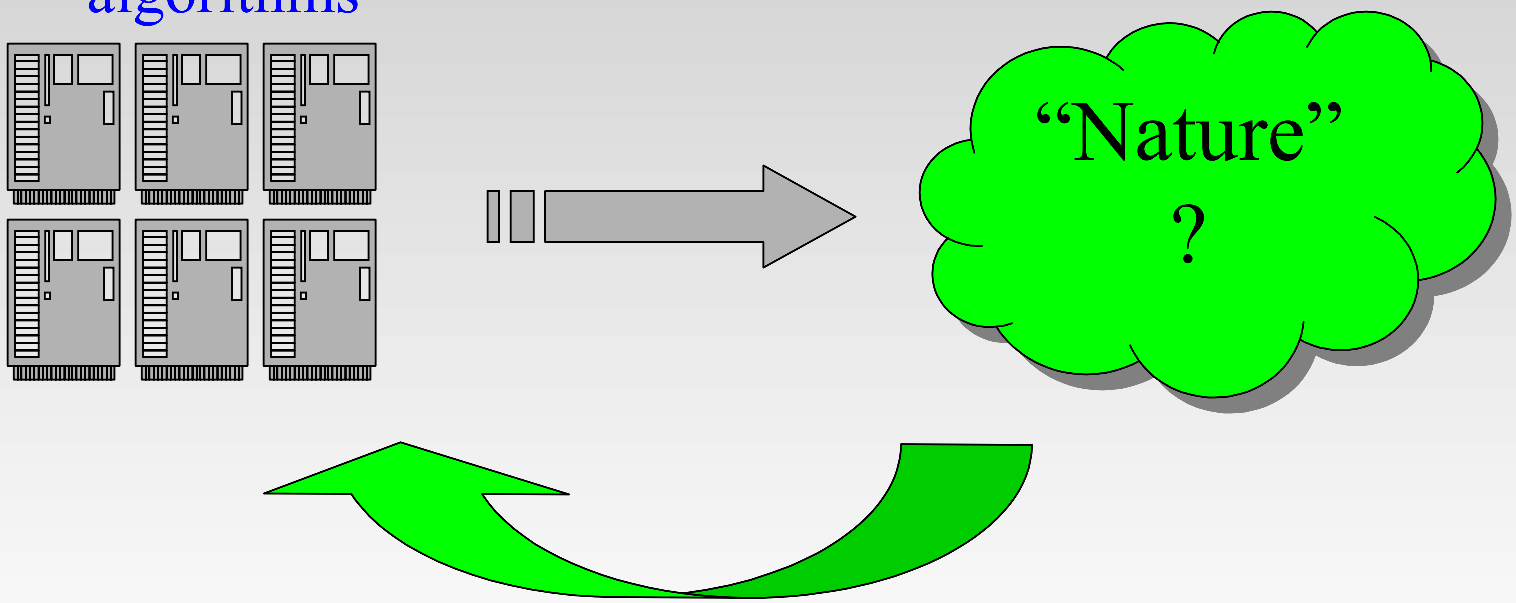
collaborators: P.A. Rikvold (Florida State U.), Z. Toroczkai (CNLS, Los Alamos), B.D. Lubachevsky, Alan Weiss (Lucent/Bell Labs)

Funded by NSF DMR/ITR,
The Research Corporation, DOE (NERSC, SCRI/CSIT, LANL),
Rensselaer, Mississippi State U.

computer architectures

+

algorithms



Discrete-event systems

- Cellular communication networks (call arrivals)
- Internet traffic routing/queueing systems

⋮

- Dynamics is asynchronous
- Updates in the local "configuration" are discrete events in continuous time (Poisson arrivals) $\Rightarrow$ discrete-event simulation

Modeling the evolution of spatially extended interacting systems: updates in "local" configuration as discrete-events

- Magnetization dynamics in condensed matter (Ising model with single-spin flip Glauber dynamic)
- Spatial epidemic models (contact process)

Parallelization for asynchronous dynamics

The paradoxical task:
- (algorithmically) parallelize (physically) non-parallel dynamics

Difficulties:
- Discrete events (updates) are not synchronized by a global clock
- Traditional algorithms appear inherently serial (e.g., Glauber attempt one site/spin update at a time)

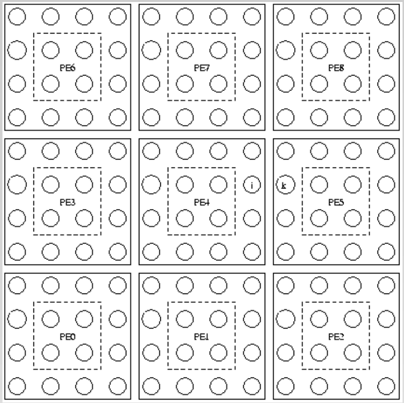
❖ However, these algorithms are not inherently serial (B.D. Lubachevsky '87)

Parallel discrete-event simulation

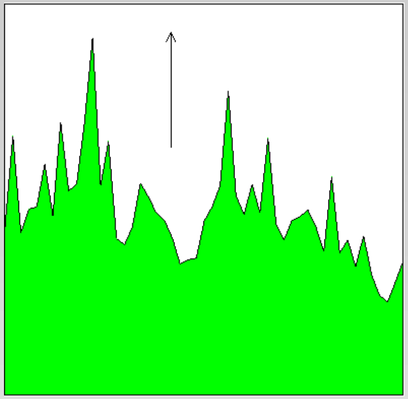
- Spatial decomposition on lattice/grid
  (for systems with short-range interactions
  only local synchronization between subsystems)
- Changes/updates: independent Poisson arrivals

❖ Each subsystem/block of sites, carried by a processing element (PE) must must have its own local simulated time, $\{\tau_i\}$ ("virtual time")
❖ Synchronization scheme
❖ PEs must concurrently advance their own Poisson streams, without violating causality

Two approaches



τ_i



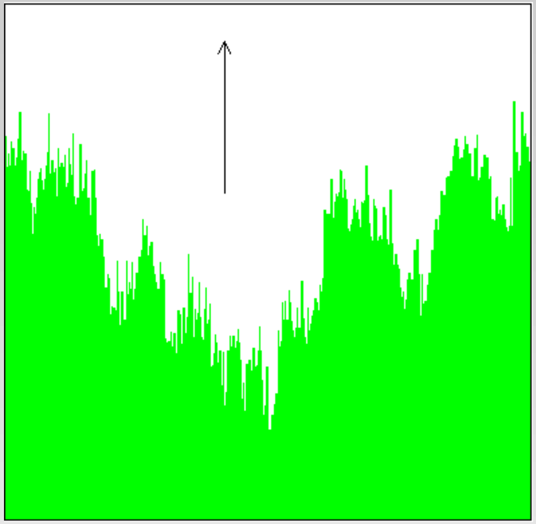
(site index) i

d=1

- ❖ Optimistic (or speculative)
  - ▪PEs assume no causality violations
  - ▪Rollbacks to previous states once causality violation is found (extensive state saving or reverse simulation)
  - ▪Rollbacks can cascade ("avalanches")

- ❖ Conservative
  - ▪PE "idles" if causality is not guaranteed
  - ▪utilization, $\langle u \rangle$: fraction of non-idling PEs

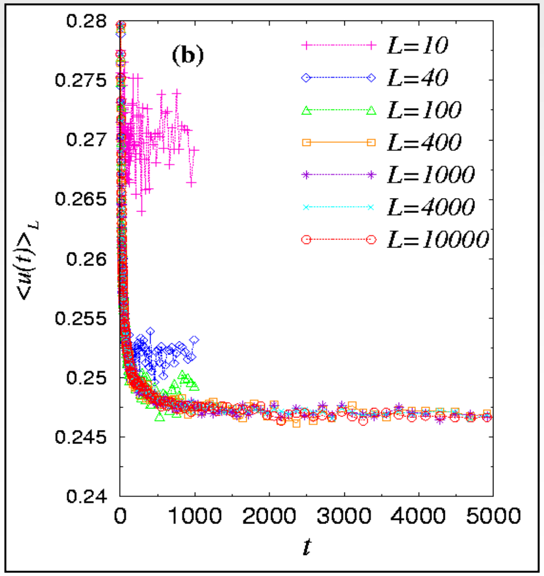
Basic conservative approach



"Worst-case" analysis:

- One-site-per PE, $N_{\mathrm{PE}}=L^d$
- t=0,1,2,… parallel steps
- $\tau_i(t)$ fluctuating time horizon
- Local time increments are iid exponential random variables
- Advance only if $\boxed{\tau_i \le \min\{\tau_{\mathrm{nn}}\}}$

(nn: nearest neighbors)



❖ Scalability modeling

- utilization (efficiency) $\langle u(t) \rangle$ (fraction of non-idling PEs) density of local minima
- width (spread) of time surface:

$$\boxed{w^2(t) = \frac{1}{N_{PE}} \sum_{i=1}^{N_{PE}} [\tau_i(t) - \overline{\tau}(t)]^2}$$

Coarse graining for the stochastic time surface evolution

G. K., Toroczkai, Novotny, Rikvold, '00

$$\tau_i(t+1)-\tau_i(t)=\Theta\big(\tau_{i-1}(t)-\tau_i(t)\big)\Theta\big(\tau_{i+1}(t)-\tau_i(t)\big)\eta_i(t)$$

•$\Theta(\ldots)$ is the Heaviside step-function
•$\eta_i(t)$ iid exponential random numbers

⋮

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

Kardar-Parisi-Zhang equation

$$P[\tau(x)]\propto\exp\left[-\frac{1}{2D}\int dx\left(\frac{\partial\tau}{\partial x}\right)^2\right]$$

Steady state (d=1): Edwards-Wilkinson Hamiltonian

❖Random-walk profile: short-range correlated local slopes

❖Universality/roughness

$$\langle w^2(t)\rangle_L \sim \begin{cases} t^{2\beta}, & \text{if} \quad t \ll t_\times \\ L^{2\alpha}, & \text{if} \quad t \gg t_\times \end{cases}$$

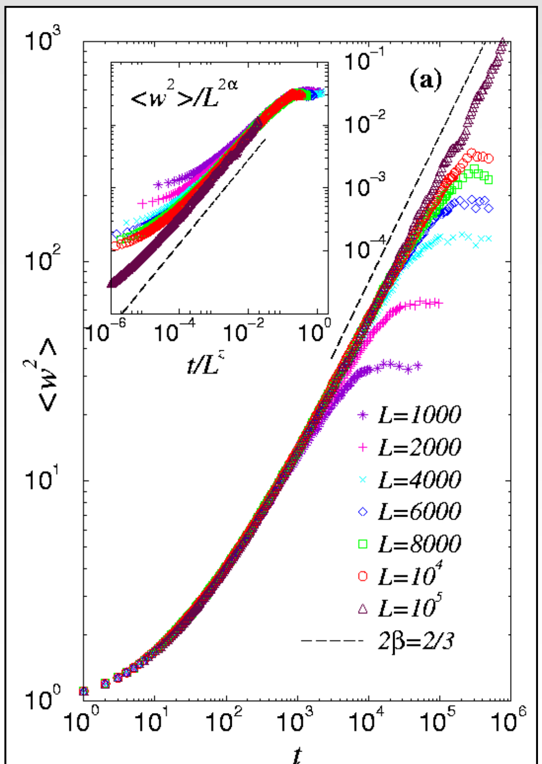
$$t \sim L^z, \quad z = \alpha/\beta$$

$$\beta \approx 0.33, \quad \alpha \approx 0.5$$

exact KPZ:

$\beta = 1/3$

$\alpha = 1/2$

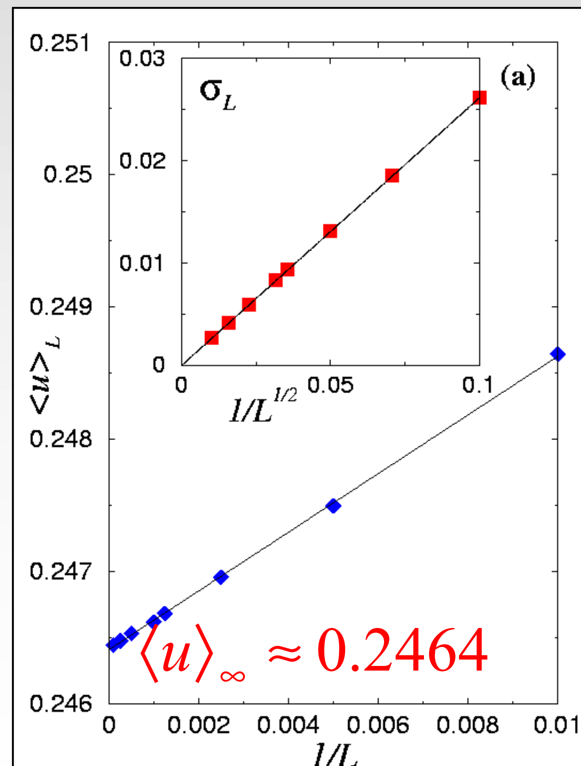


(d=1)

❖Utilization/efficiency

$$\boxed{\langle u\rangle_L \cong \langle u\rangle_\infty + \frac{const.}{L}}$$

$$\sigma_L = \sqrt{\langle u^2\rangle_L - \langle u\rangle_L^2} \sim 1/L^{1/2}$$

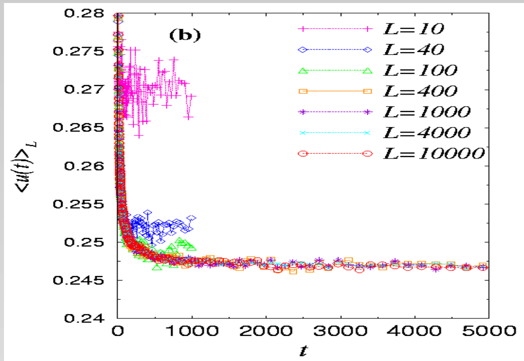


$$\langle u\rangle_\infty \approx 0.2464$$

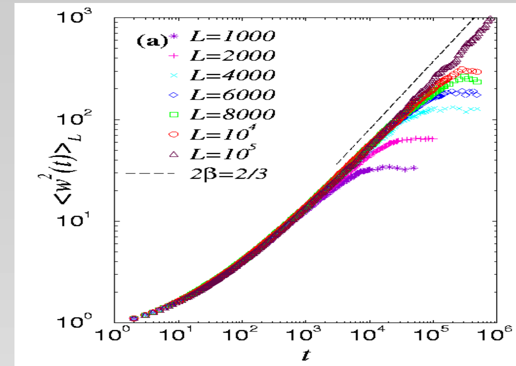
Higher-d simulations (one site per PE)

$$N_{PE} = L^d$$

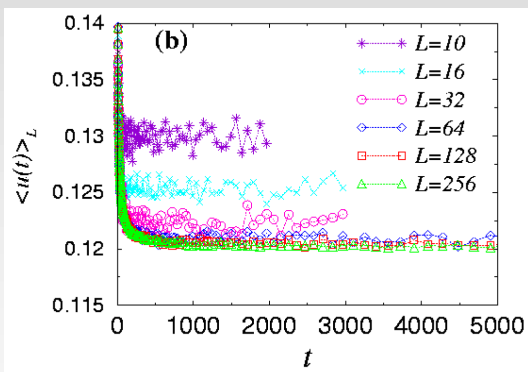
d=1



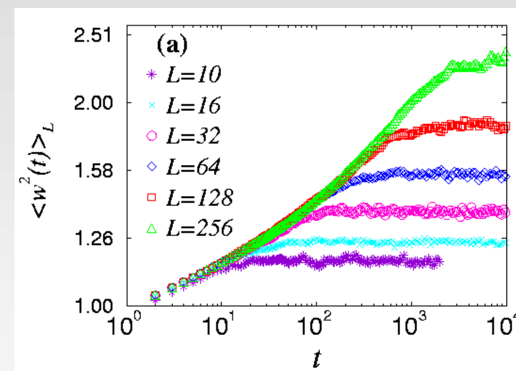
$$\langle u \rangle_\infty \approx 0.246$$



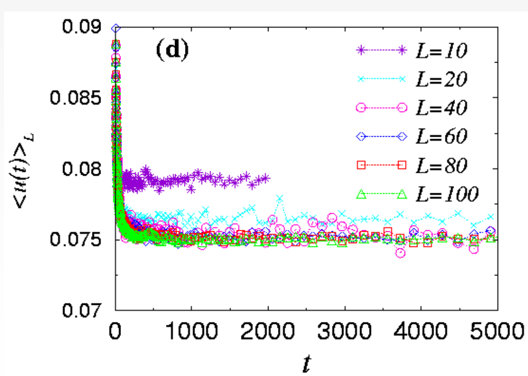
d=2



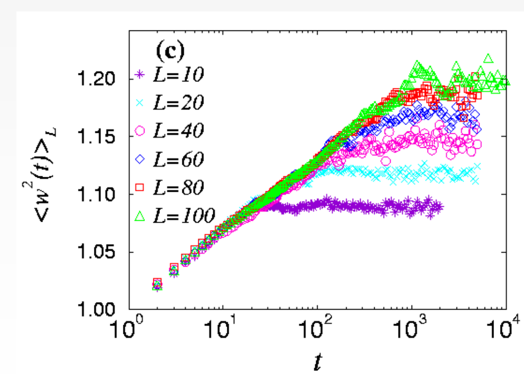
$$\langle u \rangle_\infty \approx 0.12$$



d=3



$$\langle u \rangle_\infty \approx 0.075$$



Implications for scalability

Simulation reaches steady state for $t >> L^z$ (arbitrary d)

- Simulation phase: scalable

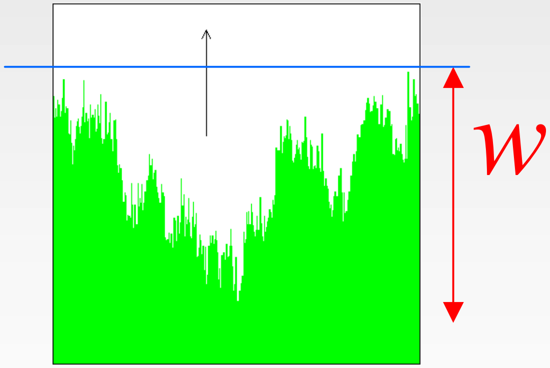
$$\langle u \rangle_L \cong \langle u \rangle_\infty + \frac{const.}{L^{2(1-\alpha)}}$$

$\langle u \rangle_\infty$ asymptotic average growth rate (simulation speed or utilization) is non-zero

Krug and Meakin, '90

- Measurement (data management) phase: not scalable

measurement at τ_{meas}: (e.g., simple averages)



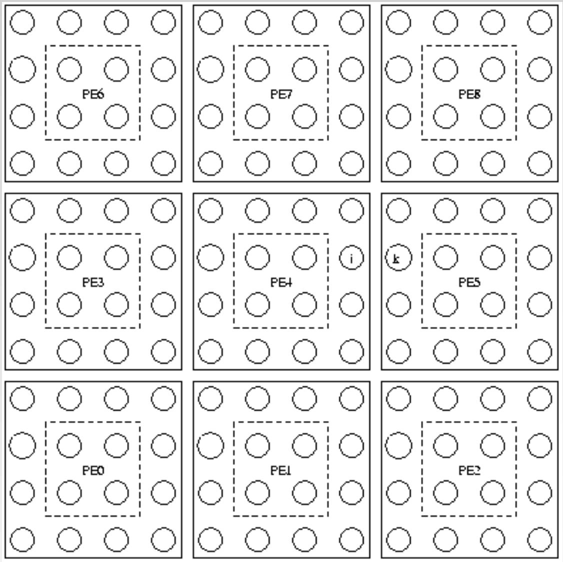
$$\langle w^2 \rangle_L \sim L^{2\alpha}$$

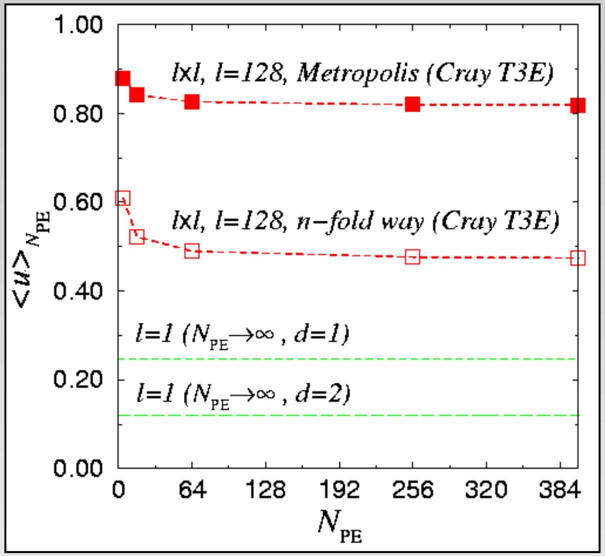
- But CAN be made scalable by considering complex underlying communication topologies among PEs

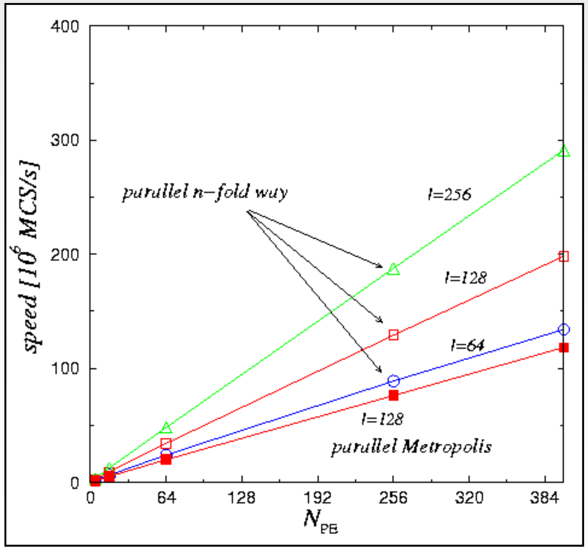
Actual implementation

$l{\times}l$ blocks

$N_{PE}=(L/l)^2$

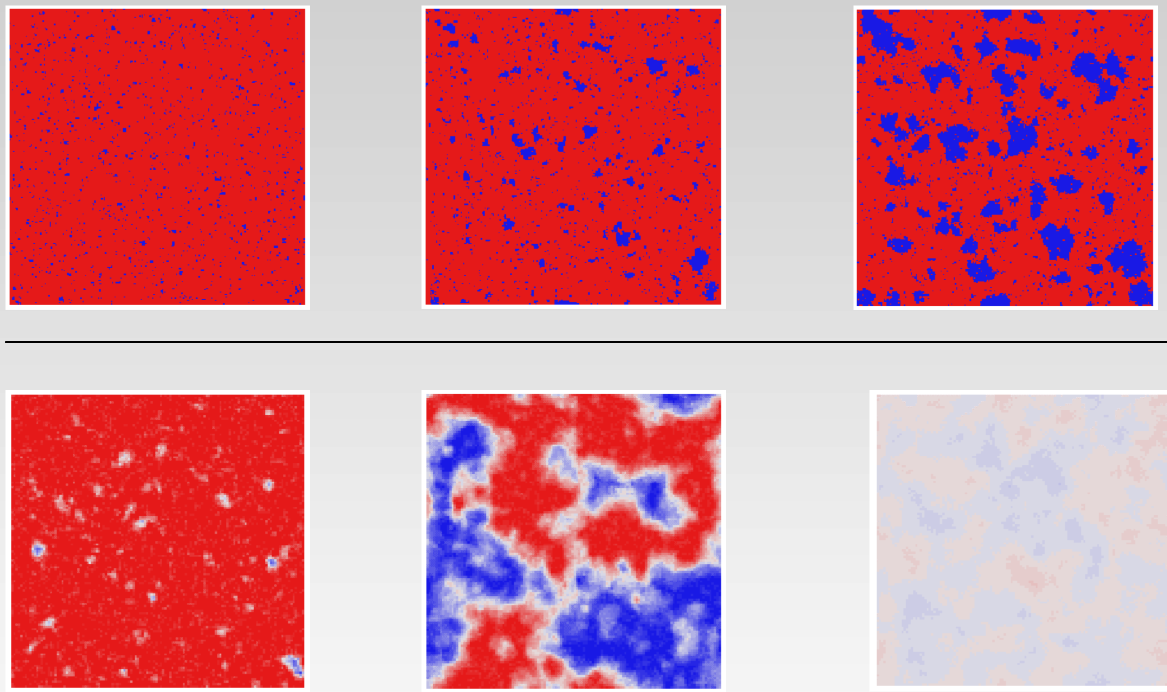






1. Local time increment:
   $\Delta\tau=-\ln(r)$, $r\in U(0,1)$
2. If chosen site is on the boundary, PE must wait until $\tau\leq\min\{\tau_{nn}\}$

Application: metastability and dynamic phase transition in spatially extended bistable systems



$\{s_i\}$

$\{Q_i\}$

$\langle \tau(T,H) \rangle$

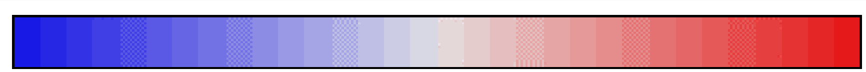
metastable lifetime

$t_{1/2}$

half-period of the oscillating field

$t_{1/2} < \langle \tau \rangle$ $t_{1/2} \approx \langle \tau \rangle$ $t_{1/2} > \langle \tau \rangle$

$$Q_i = \frac{1}{2t_{1/2}} \oint s_i(t)\,dt$$



period-averaged spin

Application: metastability and hysteresis
Kinetic Ising model

$$\mathcal{H} = -J \sum_{<i,j>} s_i s_j - H(t) \sum_{i=1}^{L^2} s_i \qquad s_i = \pm 1 \qquad J>0$$

- $L \times L$ lattice with periodic boundary conditions
- Single-spin-flip Glauber dynamics
- Periodic square-wave field of amplitude H

Half-period: $t_{1/2}$

Magnetization: $m(t)=(1/L^2)\Sigma_i s_i(t)$

$T<T_c \qquad H \rightarrow -H$

$t=0$: $m=1$ escape from metastable well: $t=\tau$: $m=0$

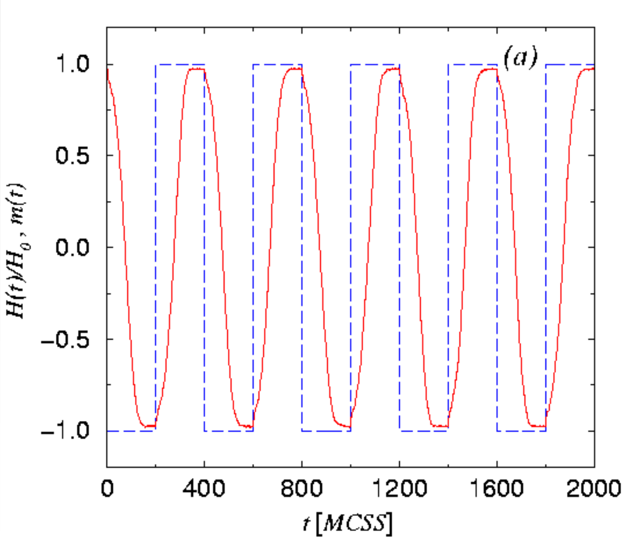
Lifetime: $\langle \tau \rangle = \langle \tau(T,H) \rangle$

Hysteresis and dynamic response

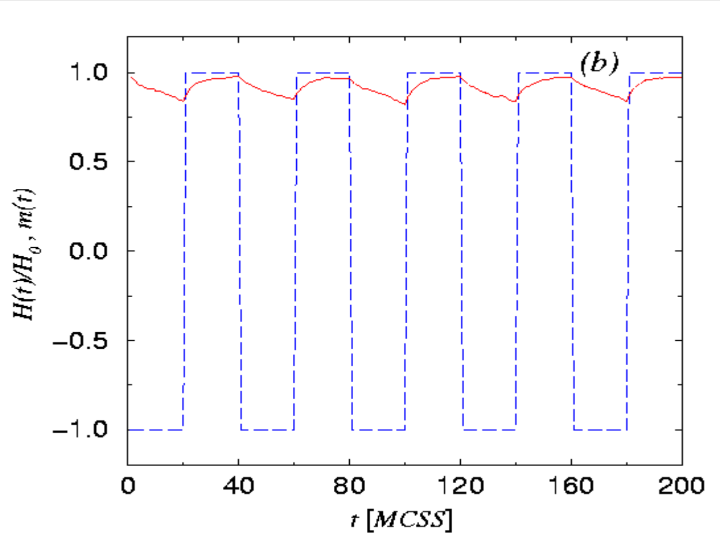
$$\mathcal{H} = -J \sum_{<i,j>} s_i s_j - H(t) \sum_i s_i$$

- Periodic square-wave field of amplitude H_o
- Half-period $t_{1/2}$; $\Theta = t_{1/2} / <\tau\,(T, H_o)>$

$\Theta >> 1$ symmetric limit cycle



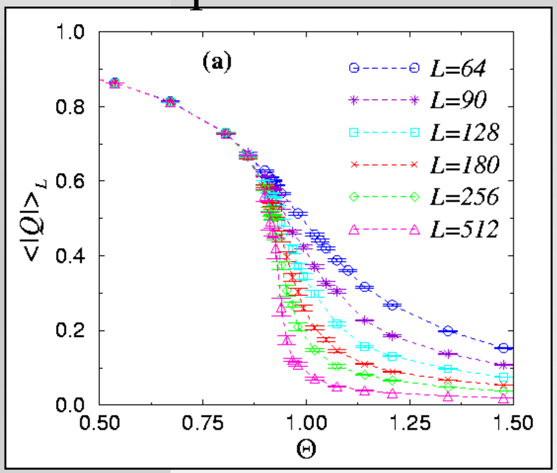
$\Theta << 1$ asymmetric limit cycle



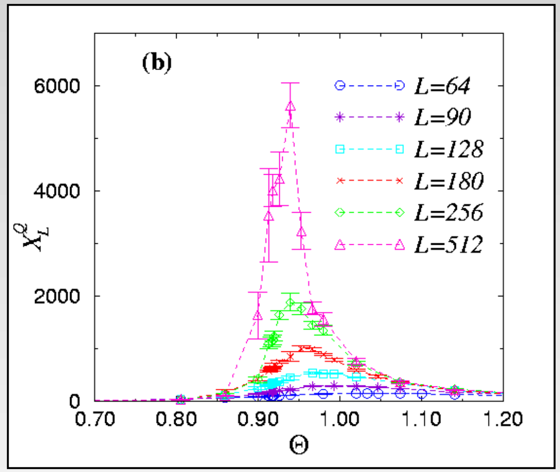
Dynamic Phase Transition (DPT)

$$Q = \frac{1}{2t_{1/2}} \oint m(t)\,dt \qquad \Theta = \frac{t_{1/2}}{\langle \tau(T,H) \rangle}$$

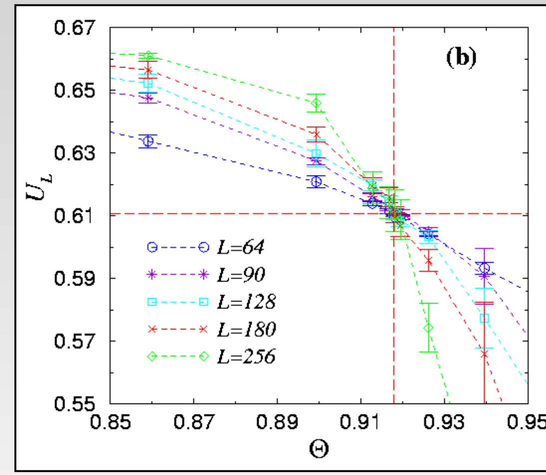
order parameter



fluctuations



4^{th} order cumulant



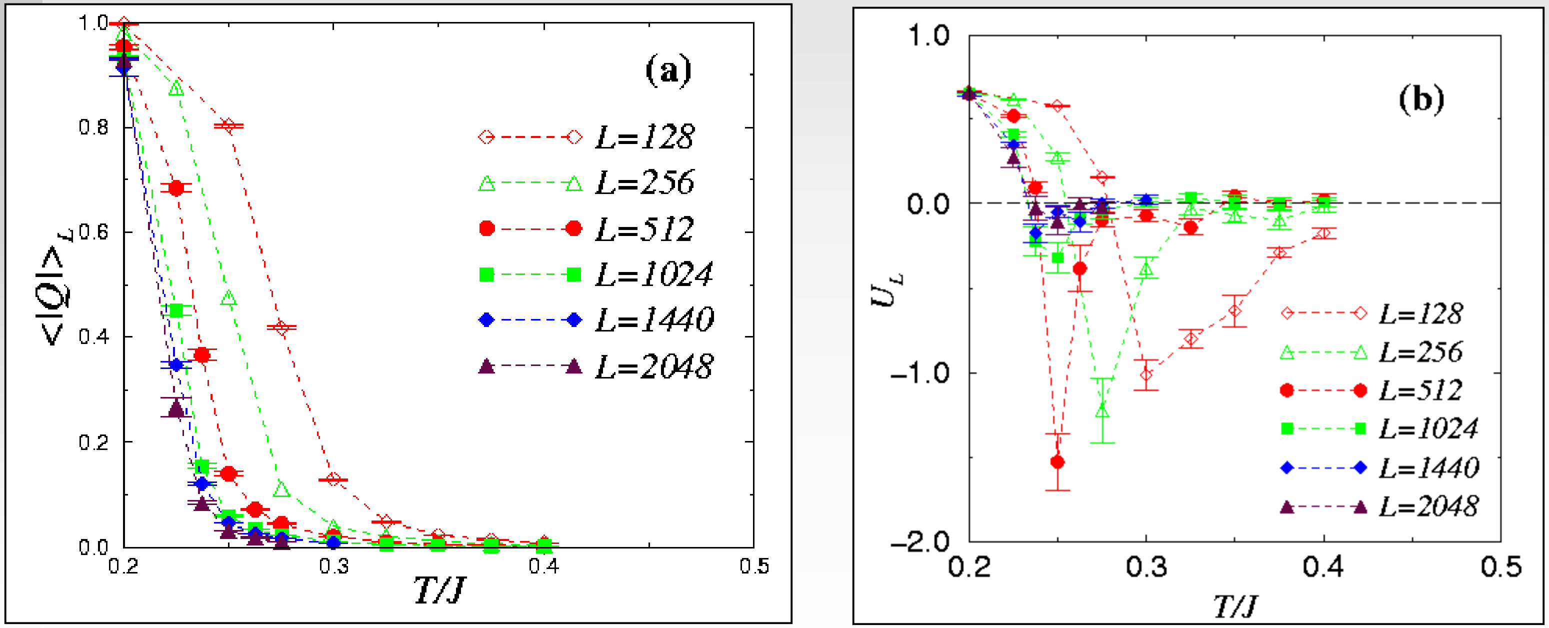
- $\Theta >> \Theta_c$: $|Q| \approx 0$ symmetric dynamic phase
- $\Theta << \Theta_c$: $|Q| \approx 1$ symmetry-broken dynamic phase
- $\Theta = \Theta_c \sim 1 \; (t_{1/2} \sim \langle \tau \rangle)$ *large fluctuations in Q* $\rightarrow$ DPT

Sides et.al., PRL’98, PRE’99
G.K. et.al., PRE’01
} finite-size scaling evidence for a continuous (dynamic) phase transition

Large-scale finite-size analysis of the dynamic phase transition : Absence of the Tri-critical Point

$$Q = \frac{1}{2t_{1/2}} \oint m(t)\,dt$$

period-averaged magnetization
(dynamic order parameter)



Summary and outlook

- The tools and machinery of non-equilibrium statistical physics (coarse-graining, finite-size scaling, universality, etc.) can be applied to scalability modeling and algorithm engineering
- Conservative schemes can be made scalable
- Optimistic schemes: rollbacks (avalanches in virtual time): Self-organized criticality ???
- Non-Poisson asynchrony (e.g., in “fat-tail” internet traffic)
- Applications: metastability, nucleation, and dynamic phase transition in spatially extended bistable systems