Complexity, Parallel Computation and Statistical Physics

Jon Machta

University of Massachusetts Amherst

Supported by the National Science Foundation

Outline

- Parallel computing and computational complexity
- Complexity of models in statistical physical
- Physical complexity and computational complexity

Complexity Journal (to appear) and cond-mat/0510809

Collaborators

- Ray Greenlaw, *Armstrong Atlantic University*
- Cris Moore, *University of New Mexico, SFI*

- Ken Moriarty, *UMass*
- Xuenan Li, *UMass*

- Dan Tillberg, *UMass*
- Ben Machta, *Brown University*

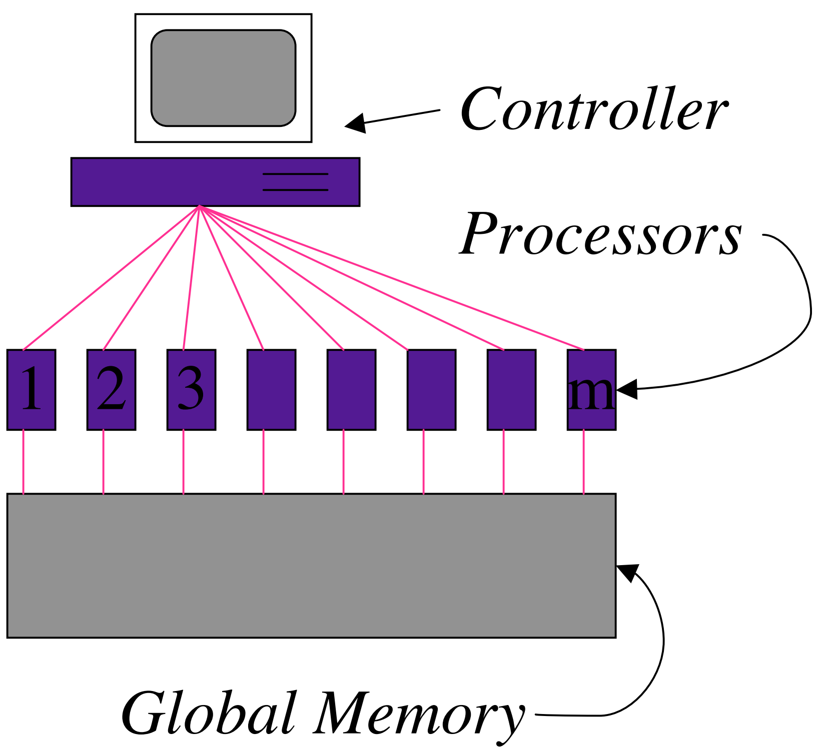
Computational Complexity

- How do computational resources scale with the size of the problem?
  - Time
  - Hardware

- Equivalent models of computation.
  - Turing machine
  - ✓**Parallel random access machine**
  - ✓**Boolean circuit family**
  - Formal logic

Parallel Random Access Machine

PRAM



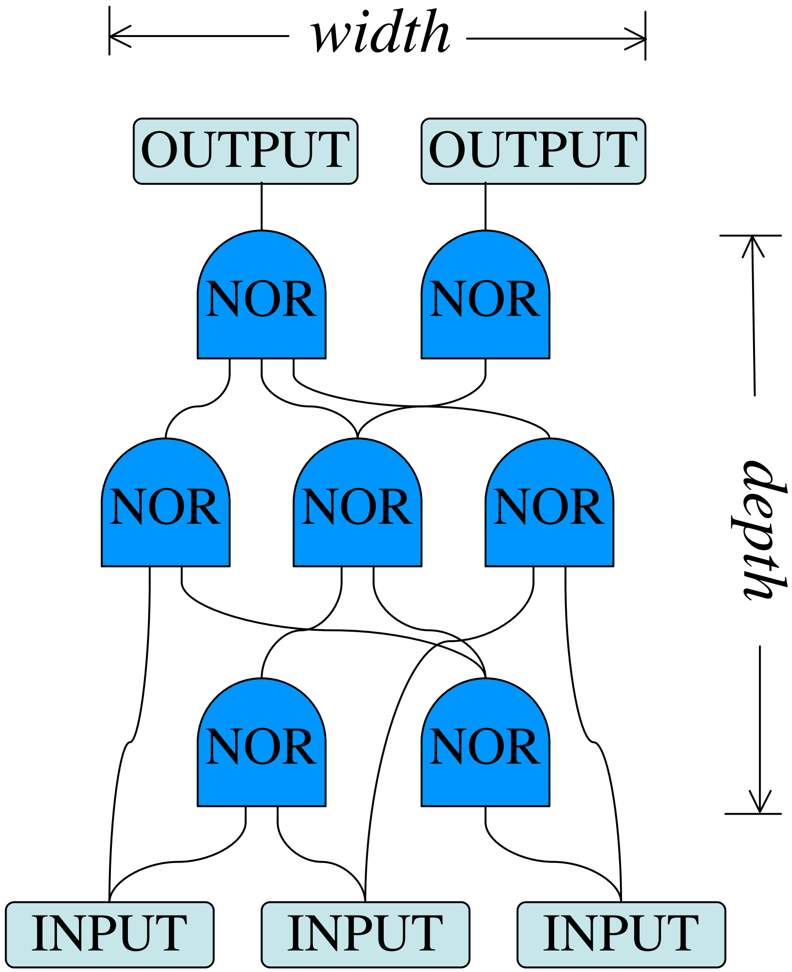
•Each processor runs the same program but has a distinct label

•Each processor communicates with any memory cell in a single time step.

•Primary resources:

- ***Parallel time***
- *Number of processors*

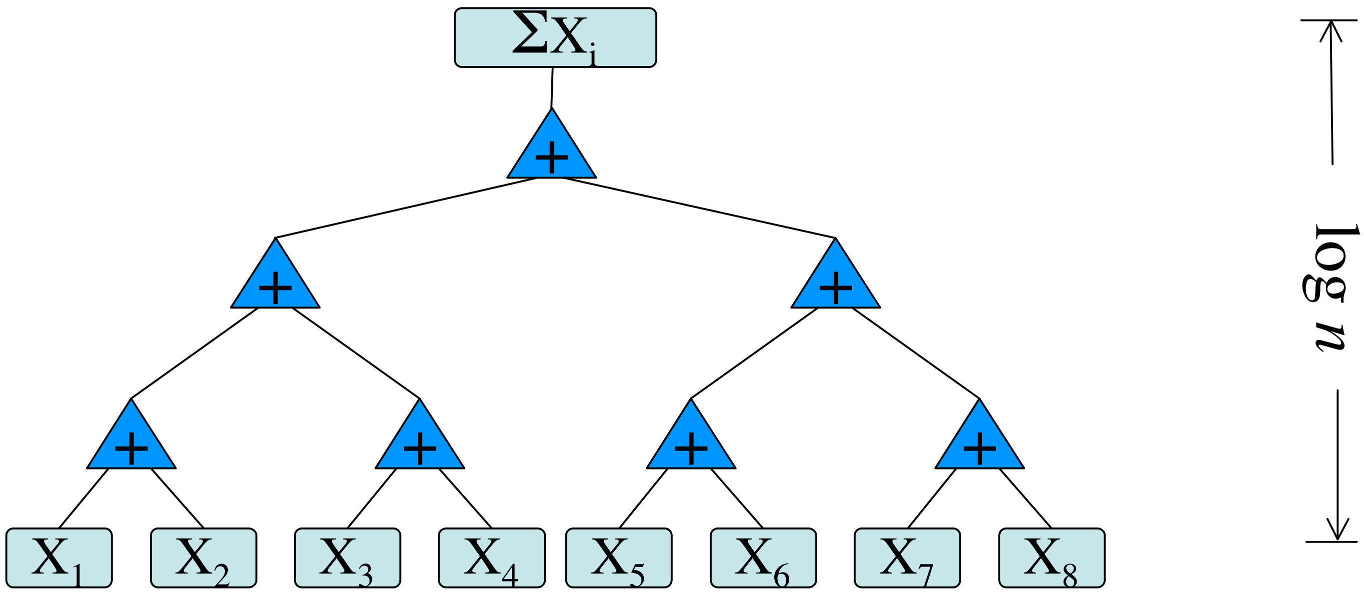
Boolean Circuit Family



- Gates evaluated one level at a time from input to output with no feedback.
- One hardwired circuit for each problem size.
- Primary resources
  - ***Depth***=number of levels
    ≈ ***parallel time***
  - *Width*=maximum number of gates in a level
    ≈*number of processors*
  - *Work*=total number of gates

Parallel Computing

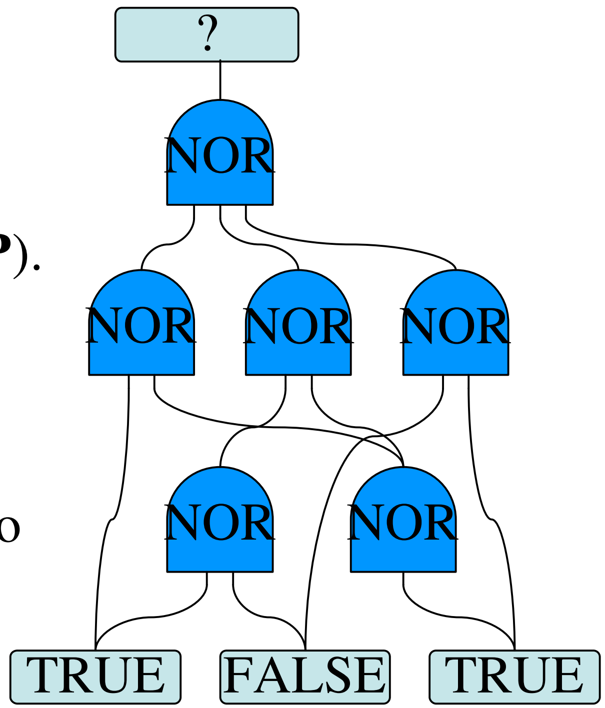
Adding n numbers can be carried out in $O(\log n)$ steps using $O(n)$ processors.



Connected components of a graph can be found in $O(\log^2 n)$ steps using n^2 processors.

Complexity Classes and P-completeness

- **P** is the class of *feasible* problems: solvable with polynomial work.
- **NC** is the class of problems efficiently solved in parallel (polylog depth and polynomial work, **NC** $\subseteq$ **P**).
- Are there feasible problems that cannot be solved efficiently in parallel (**P≠NC**)?
- **P**-complete problems are the hardest problems in **P** to solve in parallel. It is believed they are *inherently sequential:* not solvable in polylog depth.
- The Circuit Value Problem is **P**-complete.



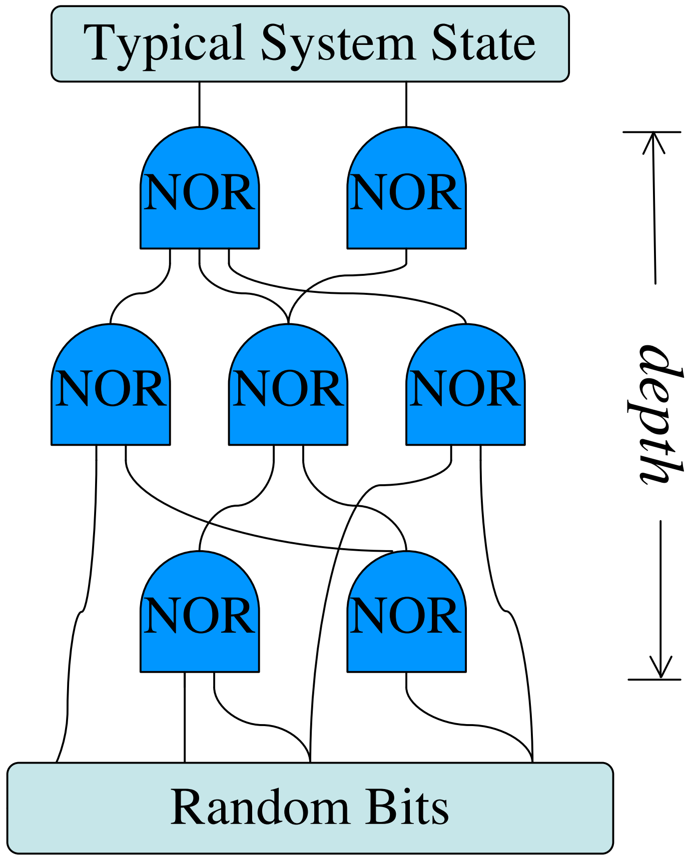
Statistical Physics

- The study of the emergent properties of many particle system using probabilistic methods.
- Objects of study are statistical ensembles of system states or histories:
  - Equilibrium states--Gibbs distribution:

$$P[\sigma] = \frac{1}{Z} \exp(-H[\sigma]/k_B T)$$

  - Non-equilibrium states--stochastic dynamics
- Computational statistical physics: sample these ensembles.

Sampling Complexity

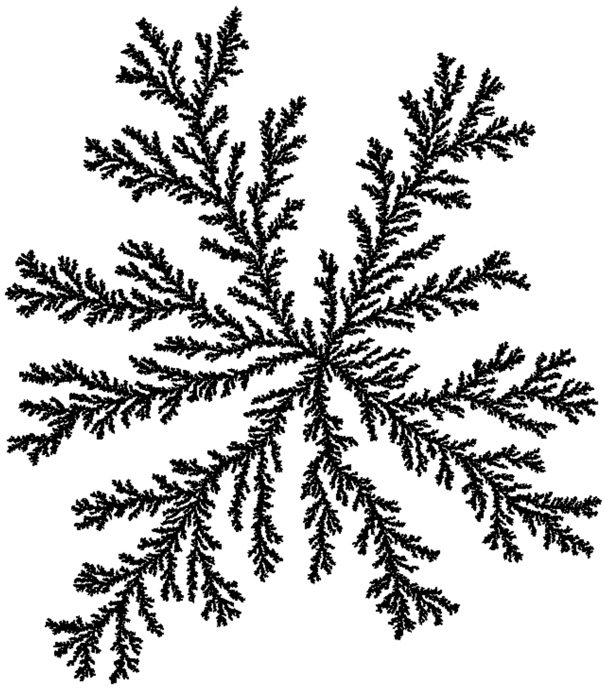


- Monte Carlo simulations convert random bits into descriptions of a typical system states.
- **What is the depth of the shallowest feasible circuit (running time of the fastest PRAM program) that generates typical states**?

Depth is a property of systems in statistical physics

Diffusion Limited Aggregation

Witten and Sander, *PRL* 47, 1400 (1981)



•Particles added *one at a time* with sticking probabilities given by the solution of Laplace's equation.

•Self-organized fractal object

d_f=1.715… (2D)

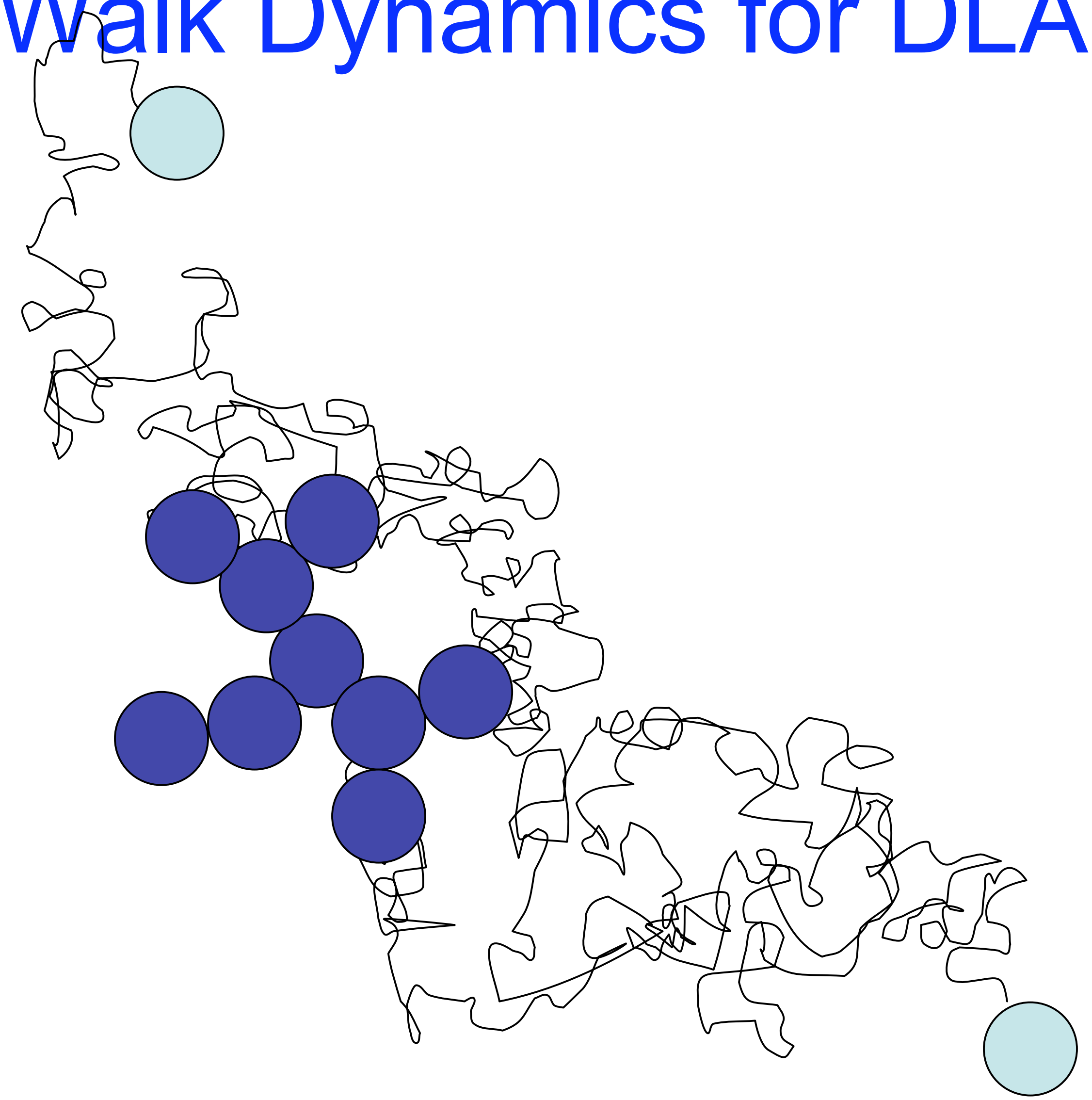
•Physical systems:

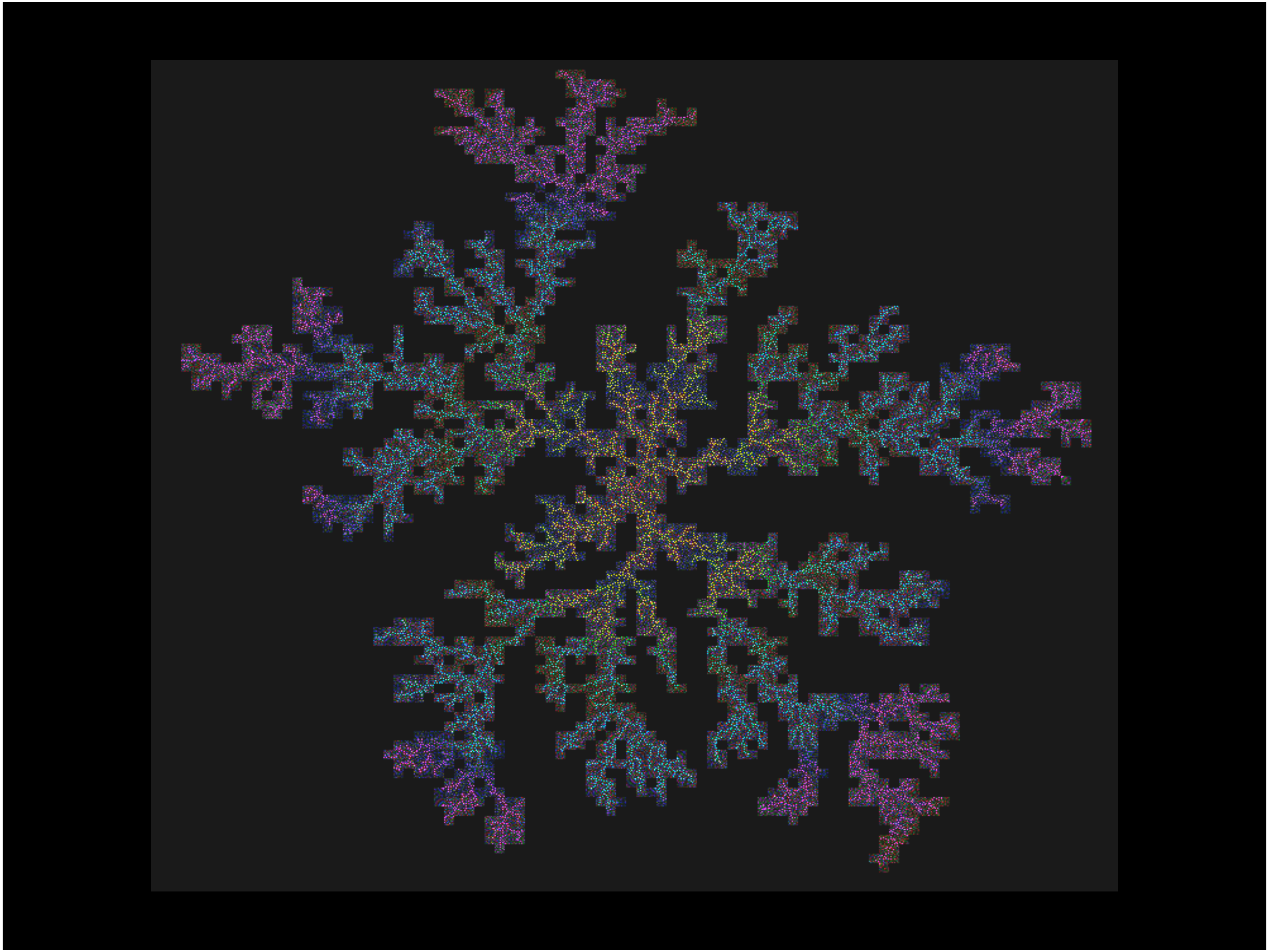
Fluid flow in porous media

Electrodeposition

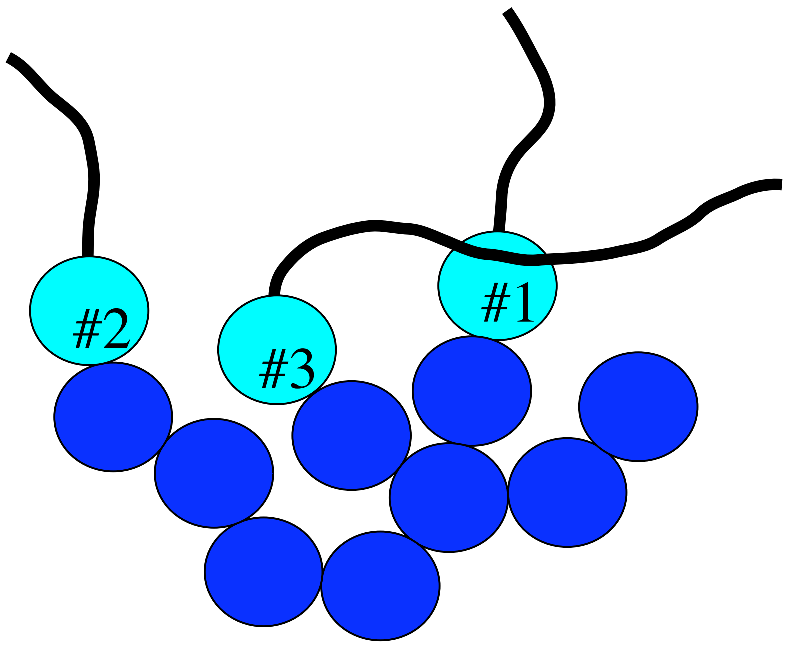
Bacterial colonies

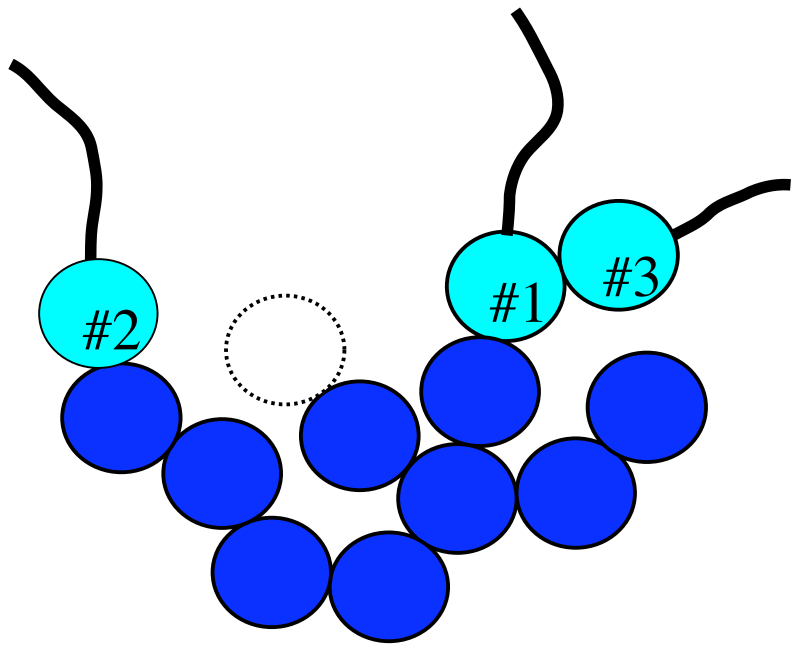
Random Walk Dynamics for DLA





The Problem with Parallelizing DLA





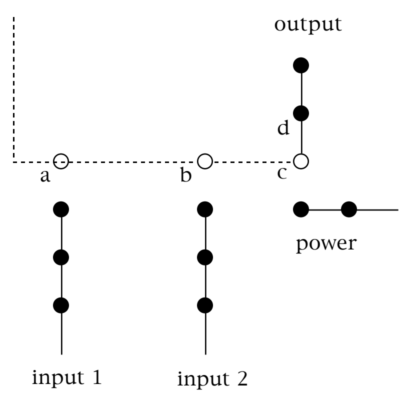
Parallel dynamics ignores *interference* between 1 and 3

Sequential dynamics

Depth of DLA

Theorem: Determining the shape of an aggregate from the random walks of the constituent particles is a **P**-complete problem.

Proof sketch: Reduce the Circuit Value Problem to DLA dynamics.



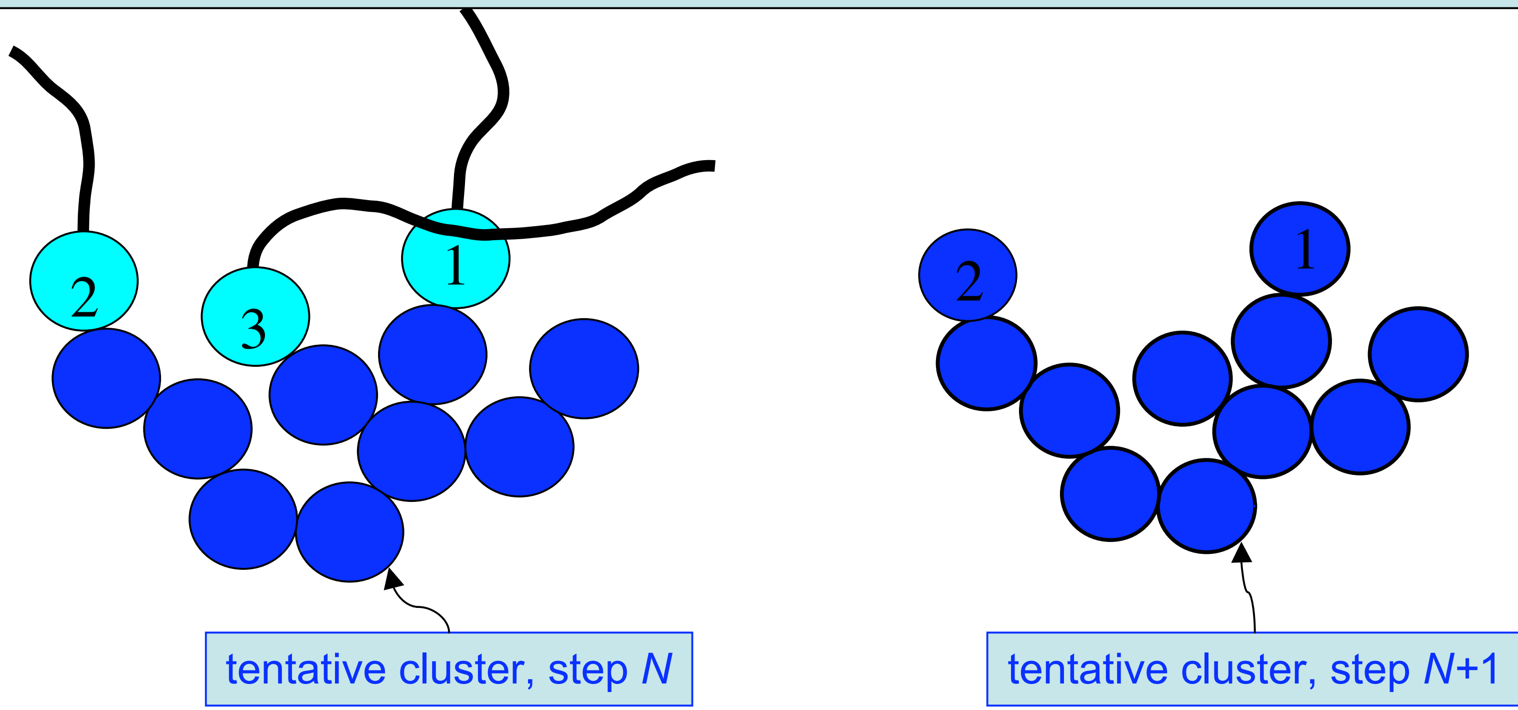
Caveats:

1. **P≠NC** not proven
2. Average case may be easier than worst case
3. Alternative dynamics may be faster than random walk dynamics

Parallel Algorithm for DLA

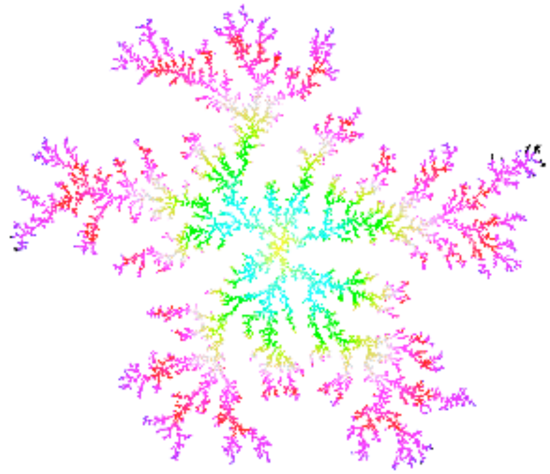
D. Tillberg and JM, *PRE* **69**, 051403 (2004)

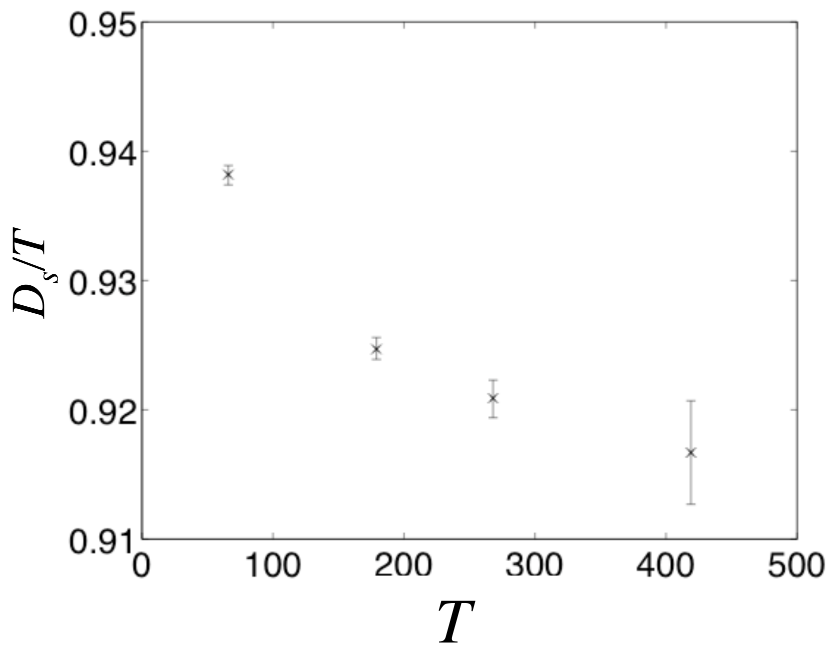
1. Start with seed particle at the origin and *N* walk trajectories
2. In parallel move all particles along their trajectories to tentative sticking points on tentative cluster, which is initially the seed particle at the origin.
3. New tentative cluster obtained by removing all particles that interfere with earlier particles.
4. Continue until all particles are correctly placed.



Efficiency of the Algorithm

- DLA is a tree whose structural depth, D_s scales as the radius of the cluster.
- The running time, T of the algorithm is asymptotically proportional to the structural depth.





$$T \sim D_s \sim N^{1/d_f}$$

Internal DLA

Particles start at the origin, random walk and stick where they first leaves the cluster.

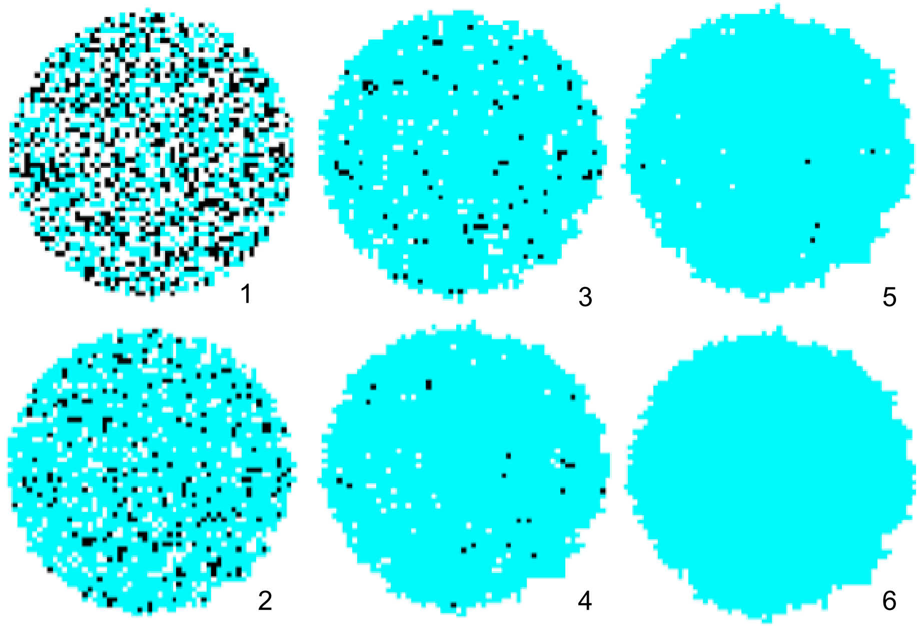
- Shape approaches a circle with logarithmic fluctuations.
- P-completeness proof fails.

Parallel Algorithm for IDLA

C. Moore and JM, *J. Stat. Phys.* **99**, 661 (2000)

1. Start with seed particle at the origin and N walk trajectories
2. Place particles at expected positions along their trajectories.
3. Iteratively move particles until holes and multiple occupancies are eliminated

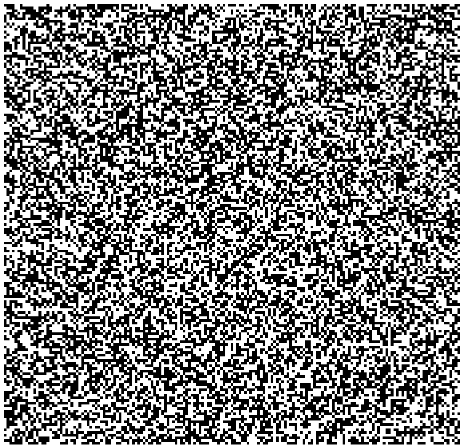
Average parallel time polylogarthmic or possibly a small power in N.



Cluster of 2500 particles made in 6 parallel steps.

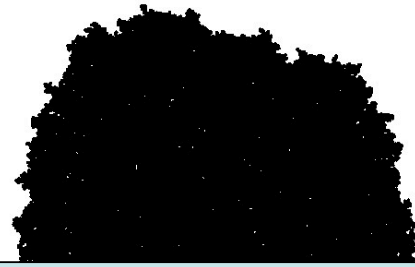
Hierarchy of Depth vs Size

constant



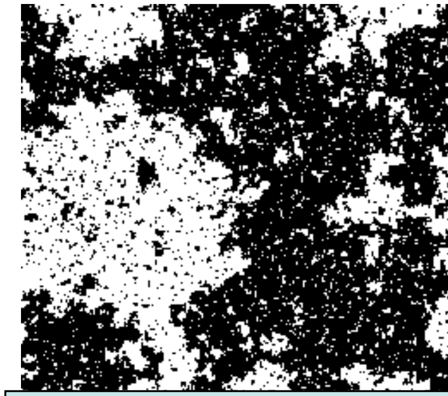
$T>T_c$ Ising

polylog

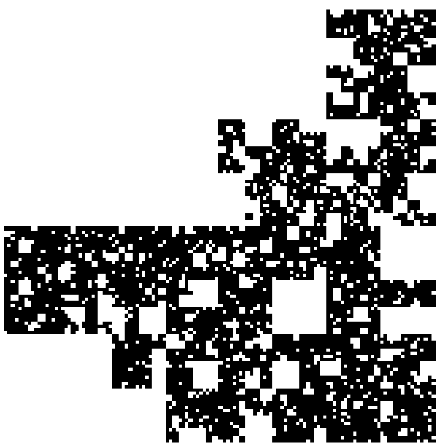


Eden growth
Invasion percolation
Scale free networks
Ballistic deposition
Bak-Sneppen model
Internal DLA

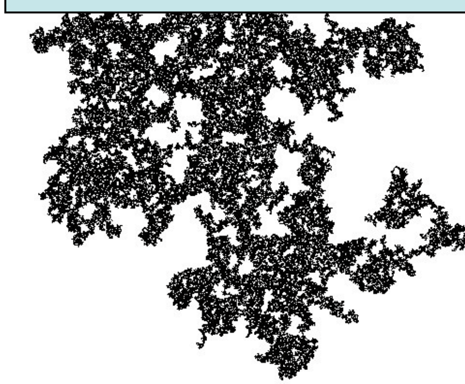
polynomial



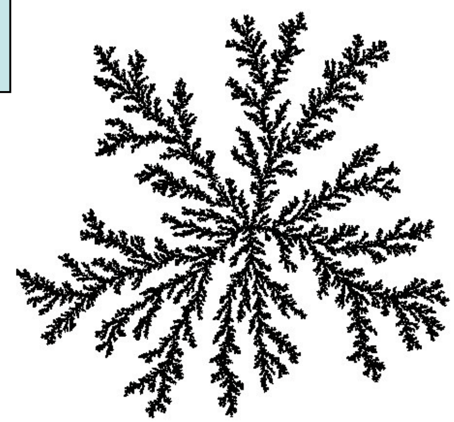
Sandpiles
Game of Life



Mandelbrot percolation



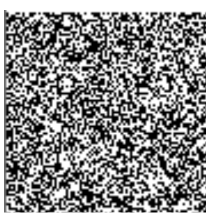
Invasion percolation



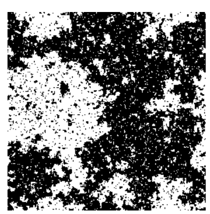
DLA

What is Physical Complexity?

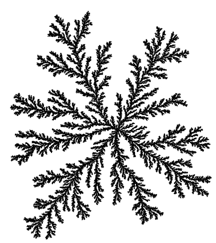
- “I shall not today attempt further to define the kinds of material I understand to be embraced with that shorthand description. ... But I know it when I see it.”
  - Justice Potter Stewart on pornography

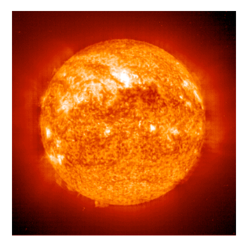


<



≤





<<

History and Complexity

–Charles Bennett

- The emergence of a complex system from simple initial conditions requires a long history.
- History can be quantified in terms of the computational complexity of simulating states of the system.

Depth is the appropriate measure to quantify history.

Depth of Physical Systems

The *depth* of a physical system is the depth of a Boolean circuit (or parallel time on a PRAM) to simulate a typical system state with polynomial hardware using the most efficient algorithm.

Maximal Property of Depth

For a system AB composed of independent subsystems A and B, the depth of the whole is the maximum over subsystems:

$$\mathcal{D}(AB) = \max\{\mathcal{D}(A), \mathcal{D}(B)\}$$

Follows immediately from parallelism.

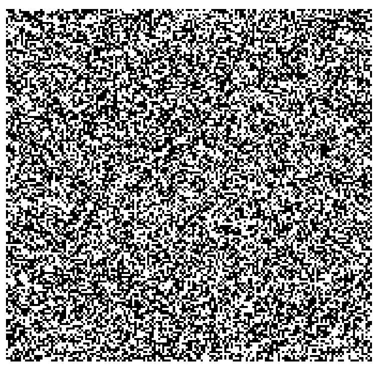
Depth is Maximal

For a system *AB* composed of independent subsystems *A* and *B*, the depth of the whole is the maximum over subsystems:

$$\mathcal{D}(AB) = \max\{\mathcal{D}(A), \mathcal{D}(B)\}$$

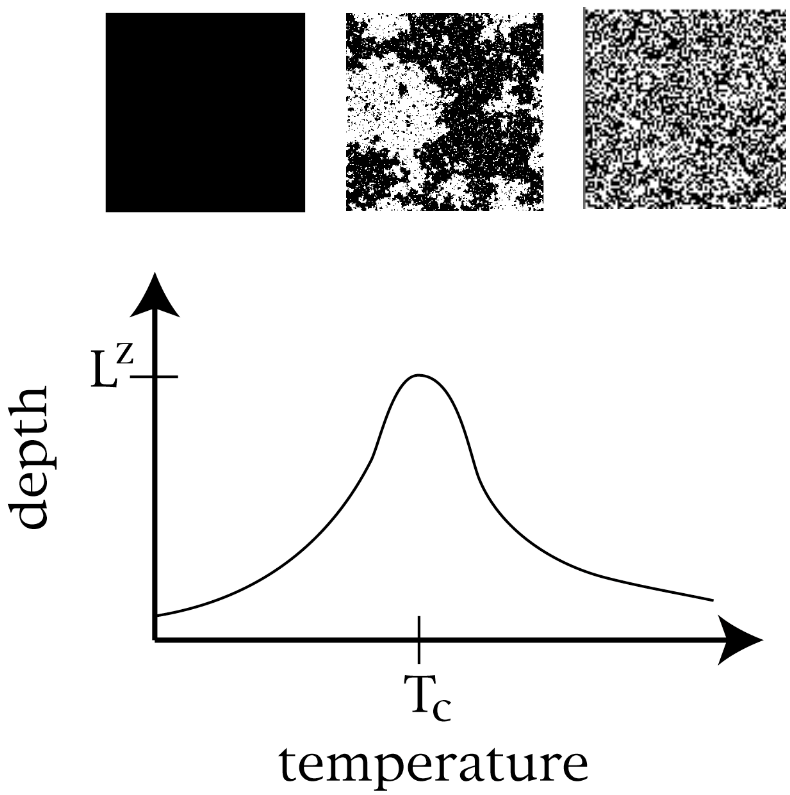
Follows immediately from parallelism.

≈



Depth is *intensive* (nearly independent of size) for homogeneous systems with short range correlations.

Depth is greatest at the boundary between order and disorder.

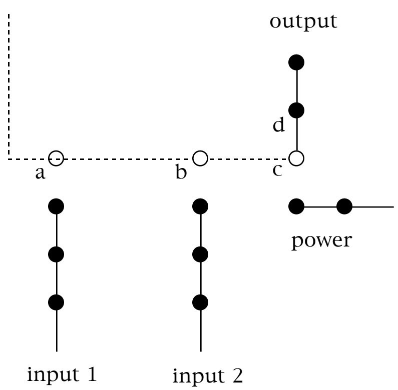


In equilibrium systems with continuous phase transitions, such as the Ising model, depth is greatest at the *critical point*, separating order and disorder (*critical slowing down*).

DLA Computes

Theorem: Determining the shape of an aggregate from the random walks of the constituent particles is a **P**-complete problem.

Proof sketch: Reduce the Circuit Value Problem to DLA dynamics.

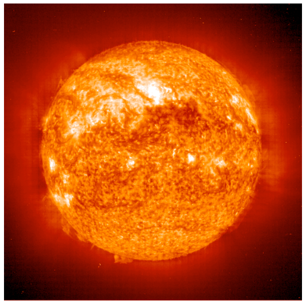


DLA dynamics can embody arbitrary computation

Properties of Depth

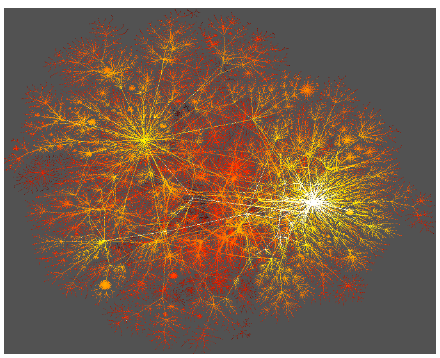
- Defined for any system in the framework of statistical physics.
- For a system of nearly independent subsystems given by the maximum over the subsystems.
- Greatest at the boundary between order and disorder.
- Systems that solve (P-hard) computational problems are deep.

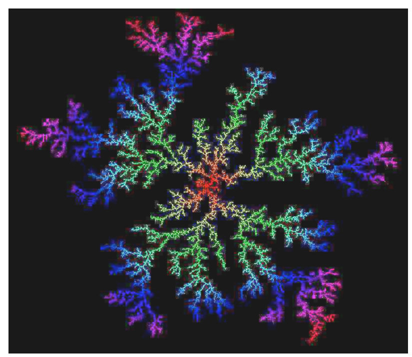
…and physical complexity



Conclusions

- Parallel computational complexity theory provides interesting perspectives on model systems in statistical physics.
- Depth, defined as the minimum number of parallel steps needed to simulate a system, is a prerequisite for physical complexity and shares many of its properties.



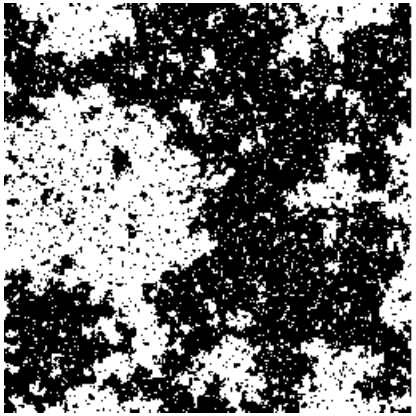


Discount Communication

 ≈ 

Complexity emerges from interactions, not from signal propagation.

Choose the Fastest Algorithm



$$\tau \sim L^z$$

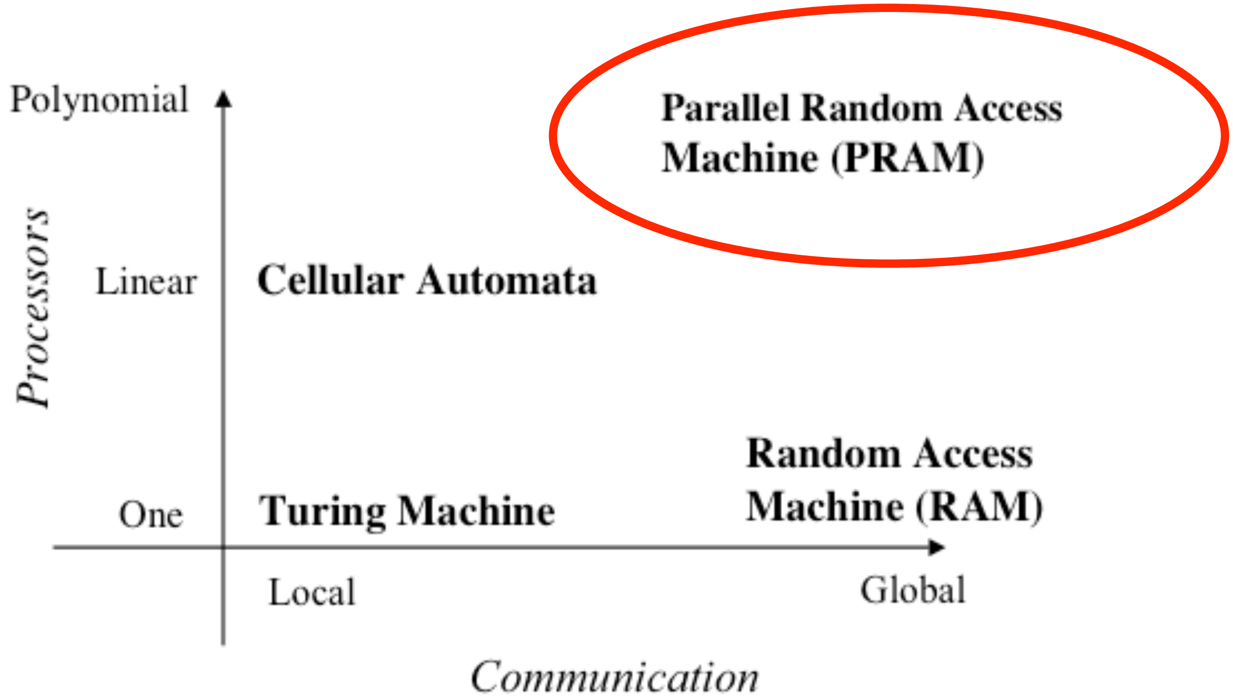
Metropolis $z \approx 2$

Swendsen-Wang $z \approx 0$

Depth of Physical Systems

The physical depth, *D(A)* of a system *A* is the average parallel time needed to generate a typical system state using the most efficient, feasible Monte Carlo algorithm for *A*

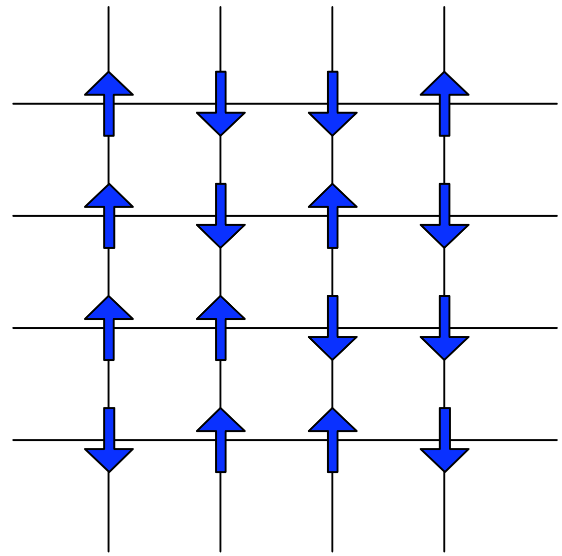
Models of Computation



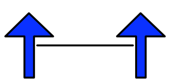
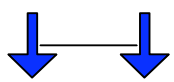
Ising Model

$$\mathcal{H} = -\sum_{<i,j>} s_i s_j$$

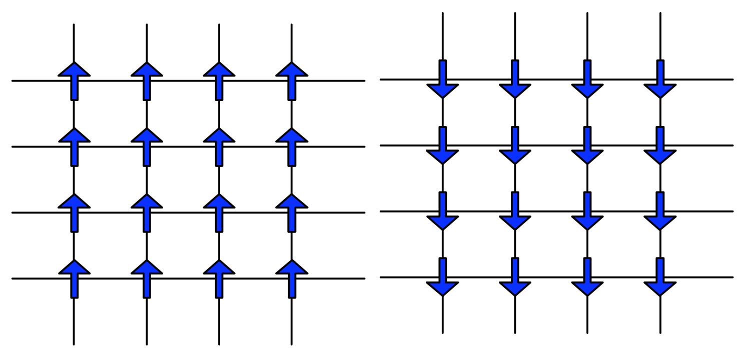
$$s_i = \pm 1$$



Aligned spins lower the energy:

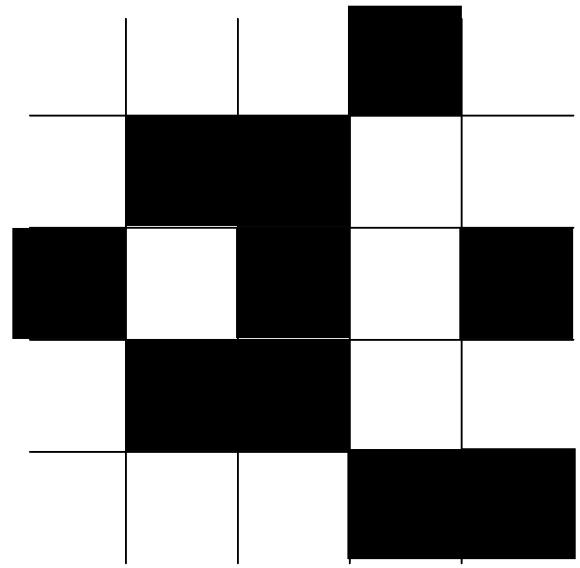
 or 

Two ground states:



Statistical Mechanics (T>0)

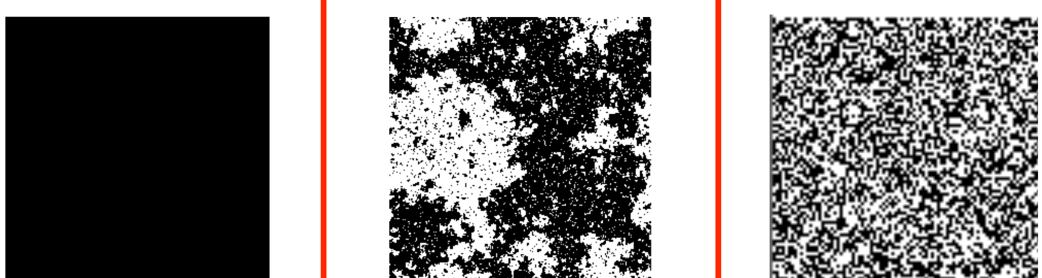
$$\mathcal{H} = -\sum_{<i,j>} s_i s_j$$



$$P[s_i] = e^{-\mathcal{H}[s_i]/kT} \Big/ Z$$

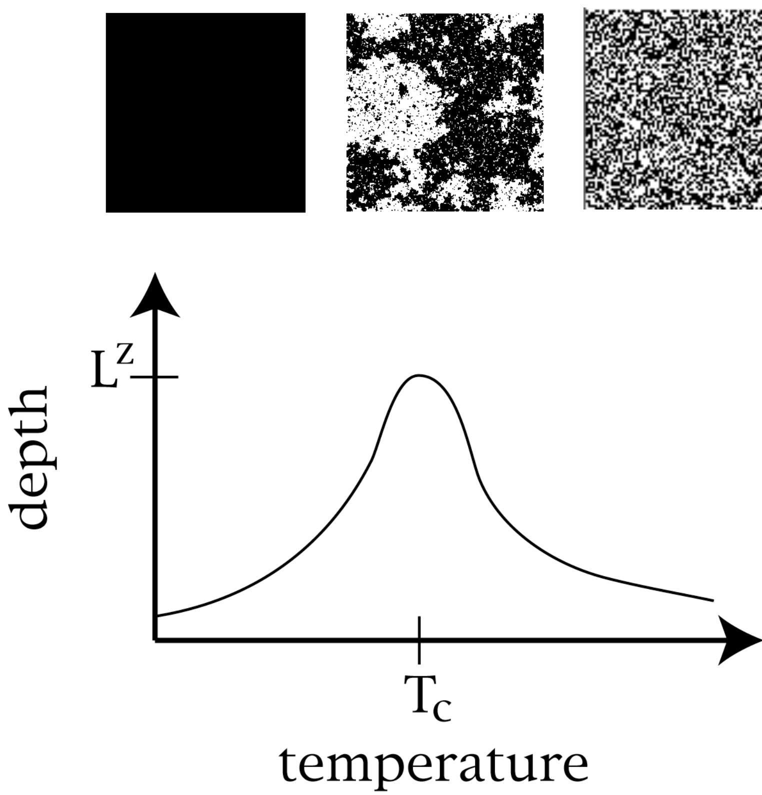
Gibbs Distribution

Critical point



T→0 $\quad$ $T = T_c$ $\quad$ T→∞

Depth of the Ising Model



Sampling time for the Ising model, is greatest at the *critical point*, separating ordered and disordered states (*critical slowing down*).

$$\tau \sim L^z$$

Metropolis $z \approx 2$

Swendsen-Wang $z \approx 0$