Large-deviation properties of random graphs

Alexander K. Hartmann
A. Engel, M. Mézard, R. Monasson

Instituts for Physics
University Oldenburg

MoN, Loughborough University, 16. September 2011

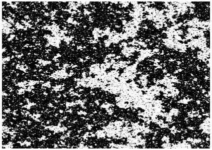
Computational Physics Group

Computer Science

Physics

helps →

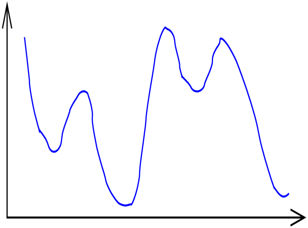
← helps



Computer simulations
new algorithms

Optimization algorithms
development/applications

[PPCC]



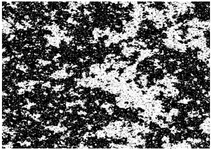
Computational Physics Group

Computer Science

helps →

← helps

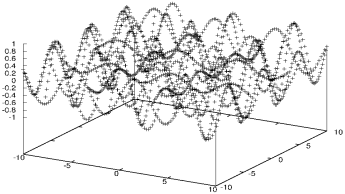
Physics



Computer simulations
new algorithms

[PPCC]

Optimization algorithms
development/applications



systems with 10^6 particles

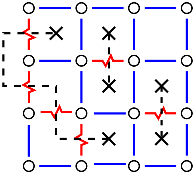
NEW:

[AKH, *Practical Guide to Computer Simulations*, World Scientific, 2009]

Disordered magnets

Spin glasses

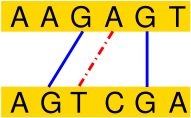
Random-field systems



Bioinformatics

RNA secondary structures

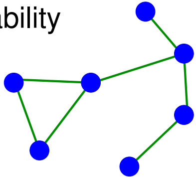
Sequence alignments



Phase transitions in optimization problems

Vertex cover

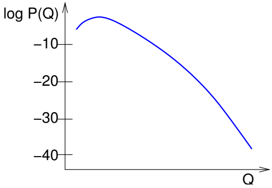
Satisfiability



Large-deviations properties

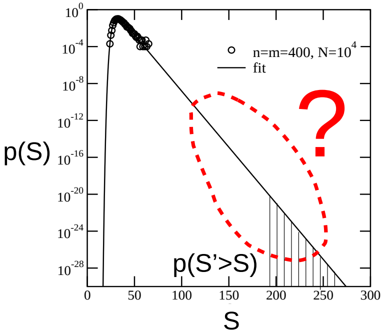
Disordered Systems

Random graphs



Large-deviation properties

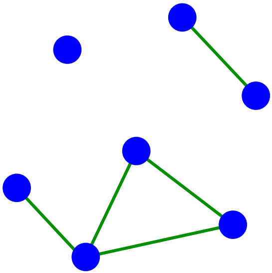
- Typical properties (probabilities $10^{-6}..1$): easy to get by simple sampling simulations
- Sometimes wanted: large deviation properties (of quenched-disorder ensembles)



- Examples:
  - Biological sequence (protein) alignment: small-probability (significant) scores [AKH, PRE 2001]
  - Distribution of the number of components of random graphs [A. Engel, R. Monasson, AKH, J. Stat. Phys. 2004]
  - Calculation of partition functions in statistical mechanics [AKH, Phys. Rev. Lett. 2005]

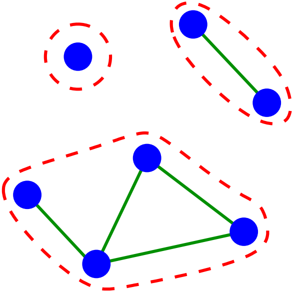
Graph ensembles

- Graph $G = (V, E)$



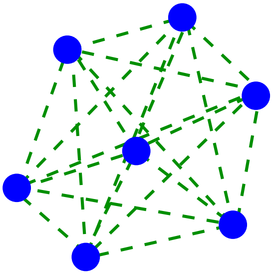
Graph ensembles

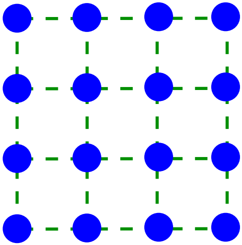
- Graph $G = (V, E)$
- connected components:
  transitive closure of “connectivity relation”



Graph ensembles

- Graph $G = (V, E)$
- connected components:
  transitive closure of "connectivity relation"
- Random graphs:
  here: N vertices, each edge tentative (ij) with prob. p.
  - Erdős-Rényi: $(ij) \in N^{(2)}$,
    $p = c/N \to$ finite connectivity c
  - two-dimensional percolation:
    $(ij) \in$ square lattice, $p =$ const



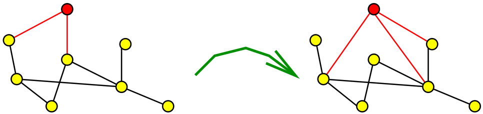


Physics Approach

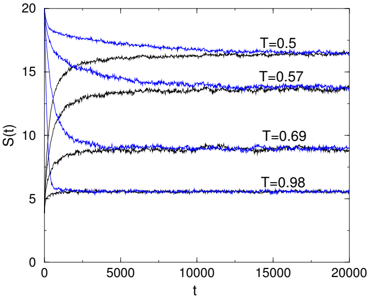
- Idea:

model	$\leftrightarrow$	physical system
quenched realisation	$\leftrightarrow$	degrees of freedom $\vec{x}$ (state)
quantity "score" S	$\leftrightarrow$	energy $E(\vec{x})$

  (ground state: often known)
  simulate at finite T
  Monte Carlo moves:
  change realisat. a bit



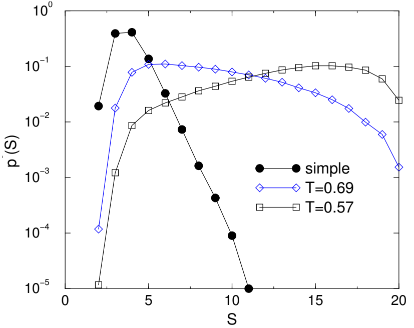
- Simulation at different T
  (using $(MC)^3$/PT)
  Example
  (sequence alignment)
  equilibration:
  start with ground state/
  with random state



- Wang-Landau approach

Distribution of Scores

- Raw result $\longrightarrow$
  (simple $\leftrightarrow T = \infty$)
  at low T:
  high scores prefered
- MC moves: $\vec{x} \to \vec{x}'$
  change on "element"
  probability $= f_a$



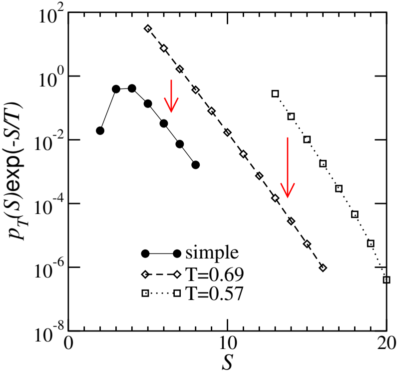
$$\Pr(\text{acceptance}) = \min\{1, \frac{\exp(S(\vec{x}')/T)}{\exp(S(\vec{x})/T)}\} = \min\{1, e^{\Delta S/T}\}$$

- $\Rightarrow$ equilibrium distribution $Q_T(\vec{x}) = P(\vec{x})e^{S(\vec{x})/T}/Z(T)$
  with $P(\vec{x}) = \prod_i f_{x_i}$, $Z(T) = \sum_{\vec{x}} P(\vec{x})e^{S(\vec{x})/T}$
  $\Rightarrow\ p_T(S) = \sum_{\vec{x},S(\vec{x})=S} Q_T(\vec{x}) = \frac{\exp(S/T)}{Z(T)} \sum_{\vec{x},S(\vec{x})=S} P(\vec{x})$
  $\Rightarrow\quad p(S) = p_T(S)Z(T)e^{-S/T}$ [AKH, PRE 2001]

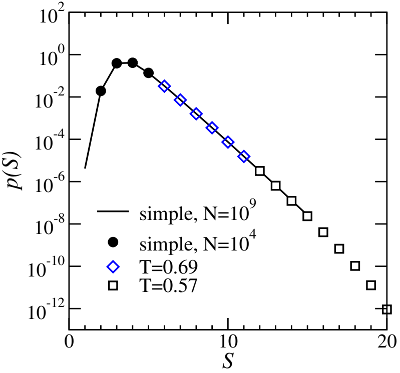
Match Distriutions

$\left[\ p(S) = p_T(S)Z(T)\exp(-S/T)\ \right]$

rescaling with $\exp(-S/T)$



$Z(T)$ by "matching"

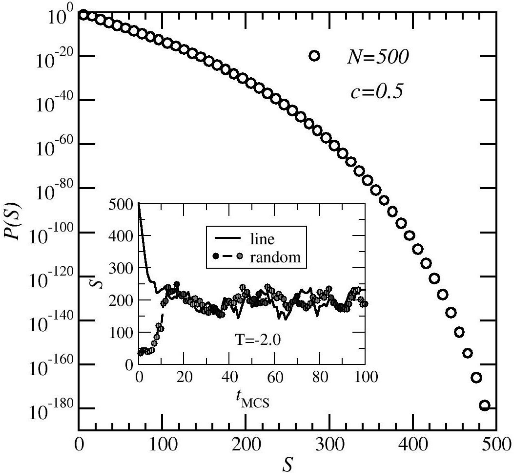


agrees with large statistics simple sampling

agrees with (for this example) known exact result

Results: Erdős-Rényi

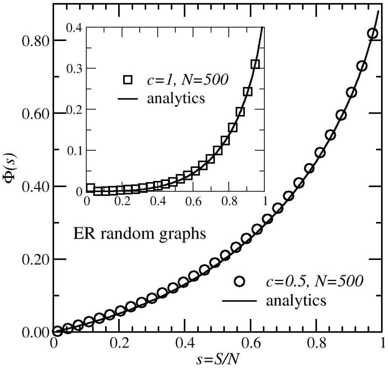
Size S of largest component (connectivity c)

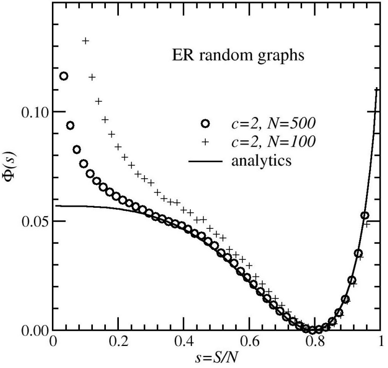


[AKH, Eur. Phys. J. B (2011)]

- Rate function $\Phi(s) \equiv -\frac{1}{N} \log P(s)$, $s = S/N$
- Comparison with exact asymptotic result

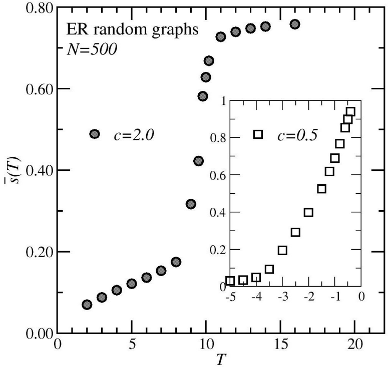
  [M. Biskup, L. Chayes, S.A. Smith, Rand. Struct. Alg. 2007]
- $\rightarrow$ evaluate algorithm $\rightarrow$ works very well
- $\rightarrow$ finite-size corrections visible

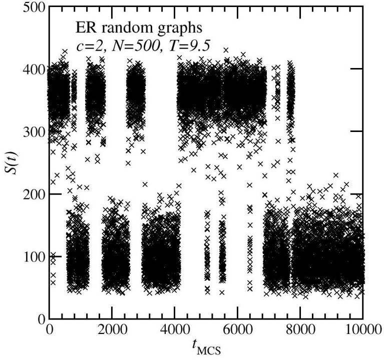




Phase transition

- Cluster size as function of (artificial) temperature
- 1st order transition in percolating phase

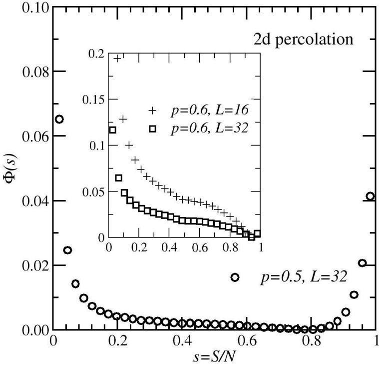


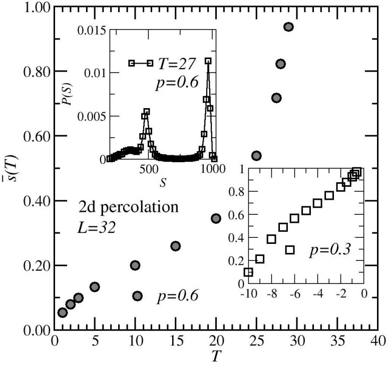


- → large system sizes not fully accessible

Two-dimensional percolation

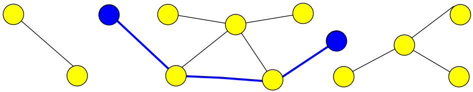
- $N = L \times L$, edge density p
- No exact result known (to me)
- Results comparable to Erdős-Rényi random graphs but stronger finite-size effects





Graph Diameter

- Diameter $d^\star$:= Longest of all shortest $i \to j$ paths



- Random graphs: ($c < 1$): Gumbel distribution

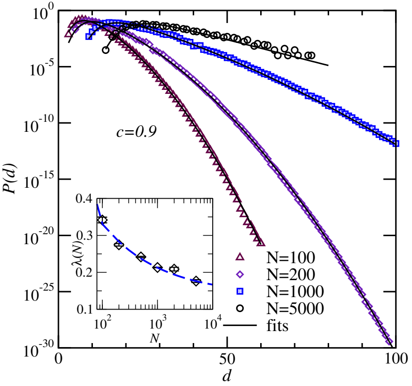
$$Pr_G(d^\star = d) = \lambda e^{-\lambda(d-d_0)} e^{-e^{-\lambda(d-d_0)}}$$

- Explanation:
  graph = forest
  d = $\max_{\text{trees } T} d(T)$
  $\to$ Gumbel distribution
- Fit to

$$P(d) = P_G(d) e^{-a(d-d_0)^2}$$

"gaussianized" Gumbel

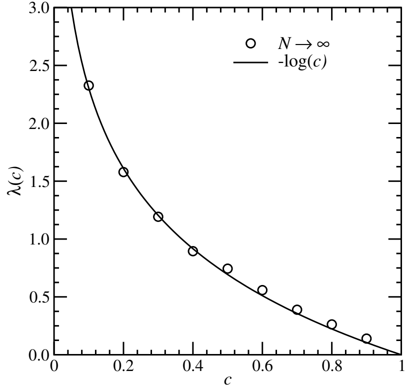
[AKH, M. Mézard, in preparation]



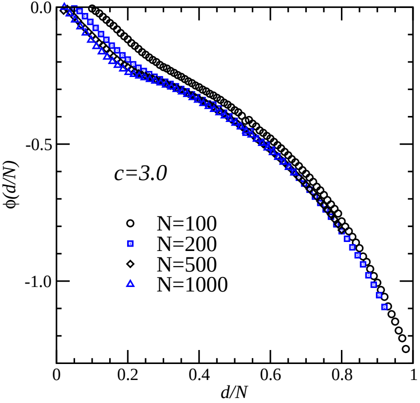
Close to $c = 1$, asymptotically

$$\lambda(c) = -\log c$$

[T. Luczak, Rand.Struct.Alg., 1998]



Percolating region:
more complex distributions



Summary

- Large-deviation properties
- Simulation approach:
  study system at artificial finite temperature
  (or, in principle, Wang-Landau algorithm $+$ modifications)
- Full distribution of size of largest component
- Erdős-Rényi random graphs: matches well analytics
  1st order transition in percolating phase
- 2d percolation: like ER model, stronger finite-size effects
- Distribution of number of components:
  agreement with statistical mechanics approach
- Distribution of diameter:
  $c < 1$: Gumbel distribution, matches theory
  $c > 1$: complex distribution, no theory