Clusters and solution landscapes for vertex-cover and SAT problems

A.K. Hartmann

Institute of Physics, University of Oldenburg

with A. Mann, W. Radenbach (Göttingen)

PhysDIS Stockholm, 17. May 2008

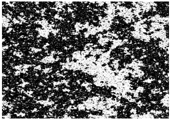
Outline

Computer Science

helps →

← helps

Physics



- Models: Vertex Cover, Satisfiability
- Analysis: neighbors-based clustering, hierarchical clustering, ballistic networking
- Results: complex/ very complex cluster structures

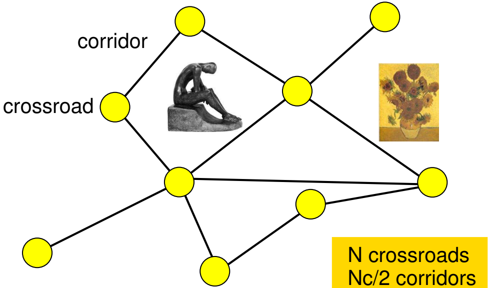
[AKH and H. Rieger, *Optimization Algorithms in Physics*, Wiley-VCH 2001]

[AKH and H. Rieger (eds.), *New Optimization Algorithms in Physics*, Wiley-VCH 2004]

[AKH and M. Weigt, *Phase Transitions in Optimization Problems*, Wiley-VCH 2005]

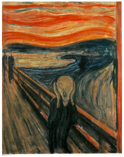
Vertex-Cover Problem (VC)

- Prototypical problem of theoretical Computer Science
- Museum

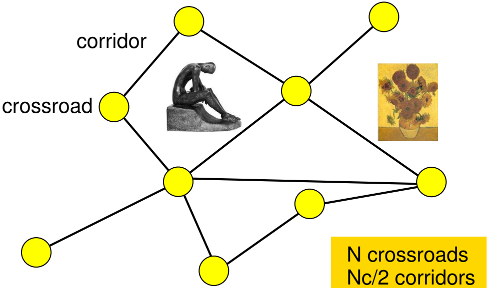


Vertex-Cover Problem (VC)

- Prototypical problem of theoretical Computer Science
- Museum
  ARE
  THEY
  SAFE?

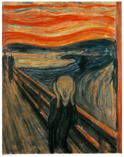


Edvard Munch's "Der Schrei" stolen in Oslo August 2004

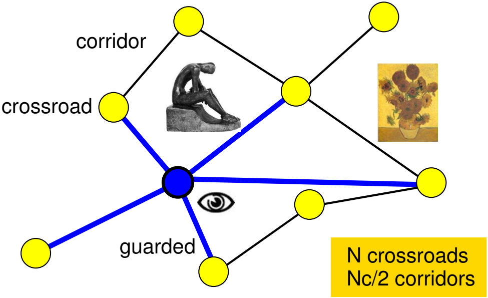


Vertex-Cover Problem (VC)

- Prototypical problem of theoretical Computer Science
- Museum ARE THEY SAFE?



Edvard Munch's "Der Schrei" stolen in Oslo August 2004

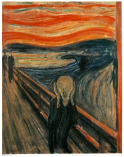


$X = xN$ guards

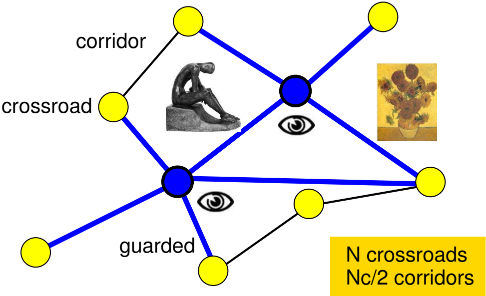
guard only adjacent corridors

Vertex-Cover Problem (VC)

- Prototypical problem of theoretical Computer Science
- Museum ARE THEY SAFE?



Edvard Munch's "Der Schrei" stolen in Oslo August 2004

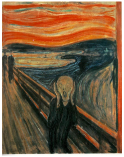


$X = xN$ guards

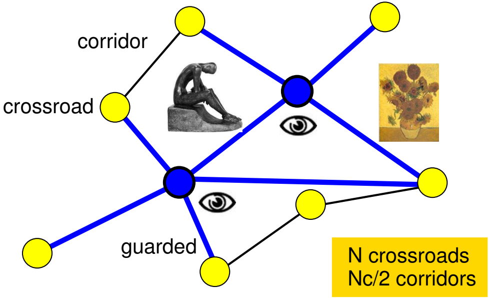
guard only adjacent corridors

Vertex-Cover Problem (VC)

- Prototypical problem of theoretical Computer Science
- Museum ARE THEY SAFE?



Edvard Munch's "Der Schrei" stolen in Oslo August 2004

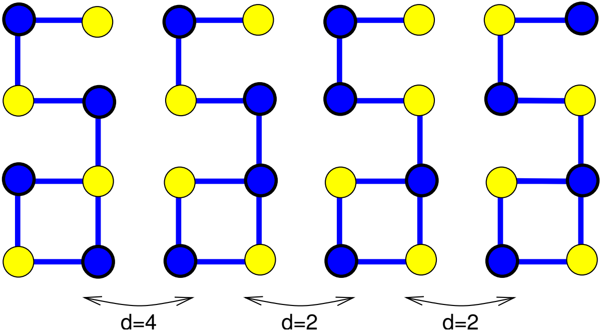


$X = xN$ guards
guard only adjacent corridors

- Mathematically: museum = graph $G = (V, E)$
  Vertex cover $A \subset V : \forall (i,j) \in E : (i \in A) \vee (j \in A)$
- Optimization problem: minimize number X of guards
- Vertex-cover problem = NP-complete
- Random graphs: phase transition at connect. $c = e \approx 2.71$

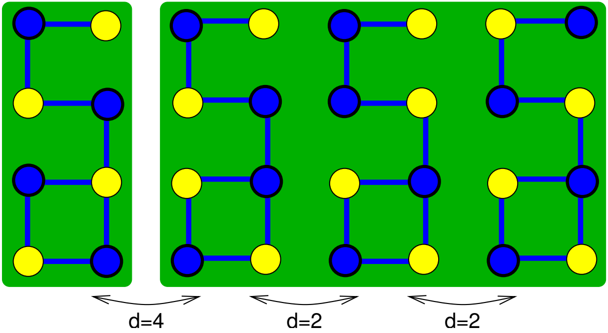
Neighbor-based Clustering

- Given: set of configurations $\{\underline{x}^i\}$
- $\underline{x}^\alpha, \underline{x}^\beta$ neigbors $\Leftrightarrow d_{\mathrm{Hamming}}(\underline{x}^\alpha, \underline{x}^\beta) \leq d_{\mathrm{max}}$.
- Example VC ($d_{\mathrm{max}} = 2$):



Neighbor-based Clustering

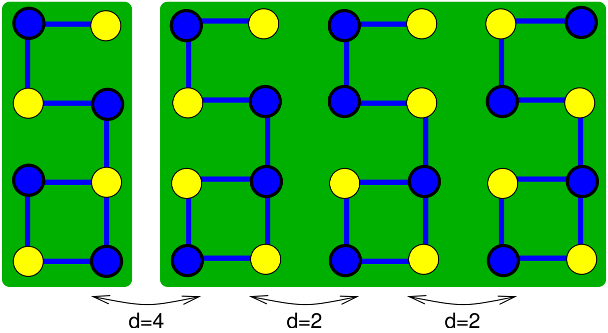
- Given: set of configurations $\{\underline{x}^i\}$
- $\underline{x}^\alpha, \underline{x}^\beta$ neigbors $\Leftrightarrow d_{\mathrm{Hamming}}(\underline{x}^\alpha, \underline{x}^\beta) \leq d_{\max}$.
- Example VC ($d_{\max} = 2$):



- Cluster: transitive closure of neighbour relation
- Algorithm: grow clusters by adding neighbors ($O(N^2)$)

Neighbor-based Clustering

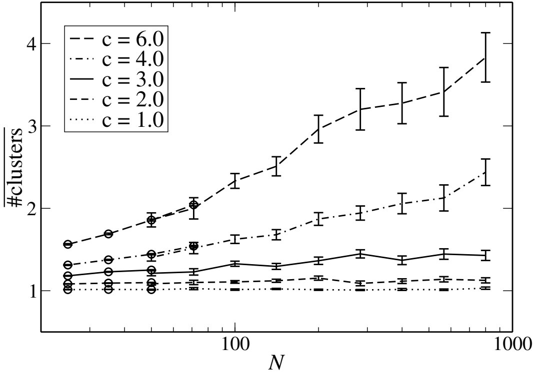
- Given: set of configurations $\{\underline{x}^i\}$
- $\underline{x}^\alpha, \underline{x}^\beta$ neigbors $\Leftrightarrow d_{\mathrm{Hamming}}(\underline{x}^\alpha, \underline{x}^\beta) \leq d_{\max}$.
- Example VC ($d_{\max} = 2$):



- Cluster: transitive closure of neighbour relation
- Algorithm: grow clusters by adding neighbors ($O(N^2)$)
- Wanted: # clusters as function of system size N

Results VC

Neighbor-based clustering ($d_{max} = 2$)



(large N: Parallel tempering (PT) and ballistic-search clustering)

$c < e$: ONE cluster, independent of N

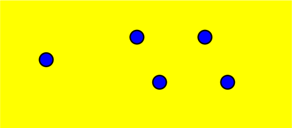
$c > e$: several clusters, logarithmic growth in N

[W. Barthel (Radenbach), AKH, Phys. Rev. E **70**, 066120 (2004)]

Hierarchical Clustering

- Start: Z configs $= Z$ single configuration clusters $C_j = \{\underline{x}^j\}$ initial distances $d(C_j, C_l) = d_{\mathrm{Hamming}}(\underline{x}^j, \underline{x}^l)$
- Merge iteratively nearest clusters $C_{\mathrm{new}} = C_\alpha \cup C_\beta$, update $d(C_{\mathrm{new}}, C_j)$ ($j \neq \alpha, \beta$), until one cluster left.

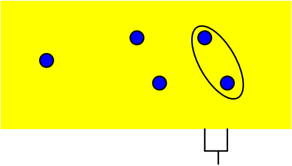
  [J.H. Ward, J. Am. Stat. Assoc. 1963]



Hierarchical Clustering

- Start: Z configs $= Z$ single configuration clusters $C_j = \{\underline{x}^j\}$ initial distances $d(C_j, C_l) = d_{\mathrm{Hamming}}(\underline{x}^j, \underline{x}^l)$
- Merge iteratively nearest clusters $C_{\mathrm{new}} = C_\alpha \cup C_\beta$, update $d(C_{\mathrm{new}}, C_j)$ $(j \neq \alpha, \beta)$, until one cluster left.

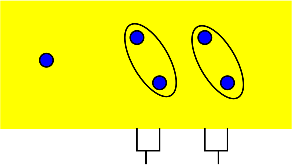
  [J.H. Ward, J. Am. Stat. Assoc. 1963]



Hierarchical Clustering

- Start: Z configs $= Z$ single configuration clusters $C_j = \{\underline{x}^j\}$ initial distances $d(C_j, C_l) = d_{\text{Hamming}}(\underline{x}^j, \underline{x}^l)$
- Merge iteratively nearest clusters $C_{\text{new}} = C_\alpha \cup C_\beta$, update $d(C_{\text{new}}, C_j)$ $(j \neq \alpha, \beta)$, until one cluster left.

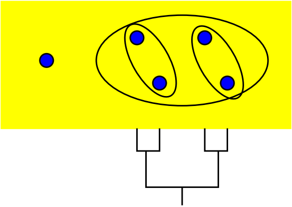
  [J.H. Ward, J. Am. Stat. Assoc. 1963]



Hierarchical Clustering

- Start: Z configs $= Z$ single configuration clusters $C_j = \{\underline{x}^j\}$ initial distances $d(C_j, C_l) = d_{\mathrm{Hamming}}(\underline{x}^j, \underline{x}^l)$
- Merge iteratively nearest clusters $C_{\mathrm{new}} = C_\alpha \cup C_\beta$, update $d(C_{\mathrm{new}}, C_j)$ ($j \neq \alpha, \beta$), until one cluster left.

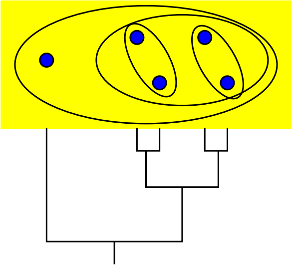
  [J.H. Ward, J. Am. Stat. Assoc. 1963]



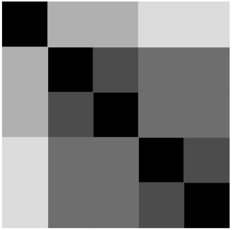
Hierarchical Clustering

- Start: Z configs $= Z$ single configuration clusters $C_j = \{\underline{x}^j\}$
  initial distances $d(C_j, C_l) = d_{\mathrm{Hamming}}(\underline{x}^j, \underline{x}^l)$
- Merge iteratively nearest clusters $C_{\mathrm{new}} = C_\alpha \cup C_\beta$, update $d(C_{\mathrm{new}}, C_j)$ ($j \neq \alpha, \beta$), until one cluster left.

[J.H. Ward, J. Am. Stat. Assoc. 1963]



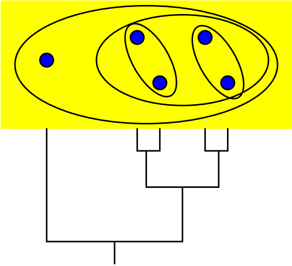
$d_{\mathrm{Ham.}}(\underline{x}^j, \underline{x}^l)$:



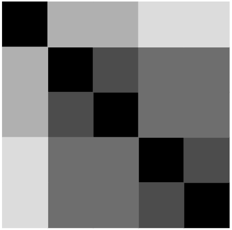
Hierarchical Clustering

- Start: Z configs $= Z$ single configuration clusters $C_j = \{\underline{x}^j\}$ initial distances $d(C_j, C_l) = d_{\text{Hamming}}(\underline{x}^j, \underline{x}^l)$
- Merge iteratively nearest clusters $C_{\text{new}} = C_\alpha \cup C_\beta$, update $d(C_{\text{new}}, C_j)$ $(j \neq \alpha, \beta)$, until one cluster left.

[J.H. Ward, J. Am. Stat. Assoc. 1963]



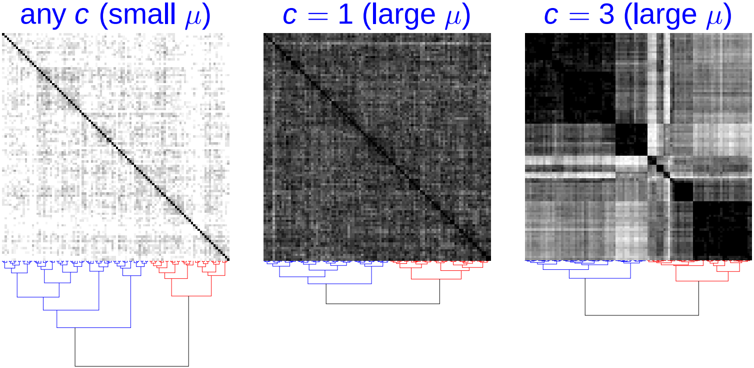
$d_{\text{Ham.}}(\underline{x}^j, \underline{x}^l)$:



- Any set of configs can be clustered → Does it match? cophenetic correlation: $\mathcal{K} \equiv [d \cdot d_c]_G - [d][d_c]_G$, (d_c: distance along tree, $[..]_G$: disorder average)

VC: hierarchical clustering

(grand-canonical ensemble (chem. pot. μ) using PT)



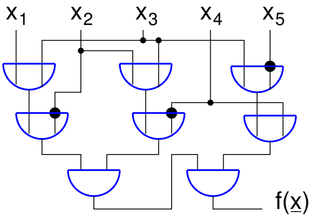
(large μ):	no structure ("paramagnet")
$c < e$:	solution cluster has no structure
$c > e$:	hierarchy of solution clusters

cophenetic correlation $K(N)$: decreases/grows for $c < e/c > e$

Complex phase space organization for $c > e$

Satisfiability

- Boolean circuit $\rightarrow$ function
- boolean variables $x_i = 0, 1$
- Operators OR, AND, NOT
- Output: $f(\underline{x}) = 0, 1$



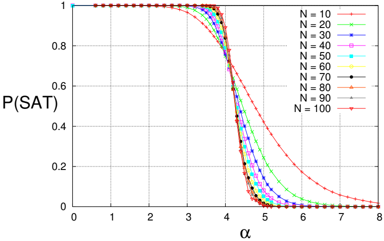
$$f(\underline{x}) = (x_1 \vee \underbrace{\overline{x_2}}_{\text{literal}} \vee x_3) \wedge \underbrace{(x_2 \vee x_3 \vee \overline{x_4})}_{\text{clause}} \wedge (x_3 \vee x_4 \vee \overline{x_5})$$

- Conjunctive normal form (CNF):
  f = conjunction of disjunctions (clauses)
- Satisfiability Problem (SAT): Is there a satisfying assignment for given f?
- SAT: “first” NP complete problem

Random 3-SAT

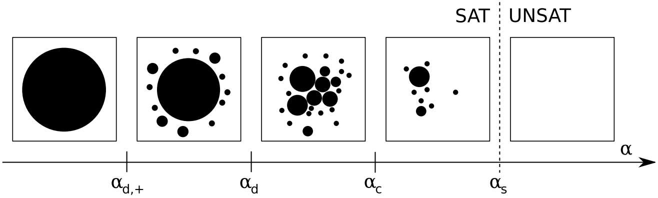
- K-SAT: CNF with K literals per clause
- N: number variables
  M: number of clauses
- 3-SAT is NP complete
- Random 3-SAT:
  literals chosen randomly.
- Phase transition in $\alpha = M/N$

  [S. Mertens, M. Mézard, R. Zecchina, RSA 2006]



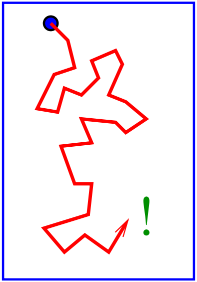
- Analytical predictions for solution-space structure

  [F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G.Semerjian, L. Zdeborová, 2007]



Average SAT (ASAT)

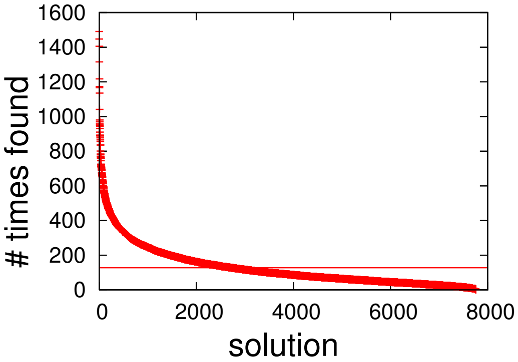
- Stochastic algorithm [J.Ardelius, E.Aurell, PRE 2006]
  $E(\underline{x}) = \#(\text{UNSAT clauses})$
  algorithm ASAT
  initialise assignment randomly
  while there are UNSAT clauses
  begin
  pick random UNSAT clause C
  pick random variable x_i from C
  if flipping x_i increases $E(\underline{x})$
  then flip x_i
  else flip x_i with prob. p
  end



- Cannot prove UNSAT.
- Solves large instances ($N = 10^6$) close to α_c ($\alpha \leq 4.2$).

Statistical Properties of ASAT

- For analyzing results:
  Each solution must contribute with same weight/probability
- How often each solution was found (sorted)
  1 realization ($N = 30$, $\alpha = 3.0$, 10^6 ASAT runs)

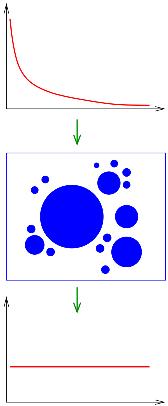


- $\Rightarrow$ ASAT is biased!

Unbias !

Outline of algorithm [AKH, EPJ B 2000], [AKH, F. Ricci-Tersenghi, PRB 2002]

1. Generate set of solutions (e.g. using ASAT)
2. Determine cluster structure $(C^1, C^2, \ldots, C^T)$
3. Determine sizes $|C^t|$ of clusters
4. Draw from each cluster C^t solutions #(solutions) $\sim |C^t|$ unbiased within cluster
5. $\Rightarrow$ unbiased set of solutions!

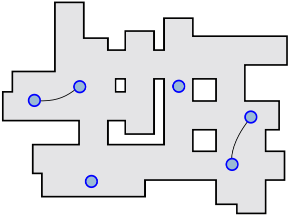


Works well up to $N = 256$.

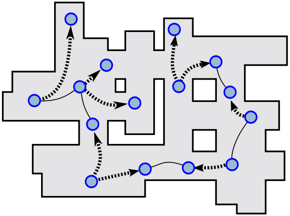
Details follow now!

Ballistic search

- Cluster := neighbor-based clusters with $d_{\text{max}} = 1$
- Two solutions $\underline{x}^\alpha$,$\underline{x}^\beta$ in same cluster $\Leftrightarrow$ connected by zero-energy single-variable flips
- Ballistic search:
  1. Start at $\underline{x}^\alpha$.
  2. Flip zero-energy variables with $x_i^\alpha \neq x_i^\beta$, each at most once
  3. If arrival at $\underline{x}^\beta \rightarrow$ same clusters else ??



- Improvement: Ballistic networking:
  Per representing solution:
  Generate additional ($N_{\text{add}} = 5$) attached solutions via $T = 0$ MC.
  Perform ballistic search pairwise for attached solutions.



- Iterate until cluster structure is stable.

Cluster Size

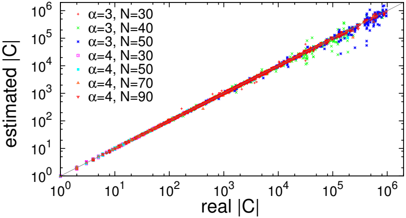
- Given: any solution $\underline{x}^0$ from cluster C
- Test Hamiltonian [AKH, F. Ricci-Tersenghi, PRB 2002]

$$H_{\text{test}}(\underline{x}) = d_{\text{Hamming}}(\underline{x}, \underline{x}^0) \quad (\underline{x} \in C)$$

$S(\beta)$=entropy $\Rightarrow |C| = \exp(S(\beta = 0))$
From thermodynamic integration

$$S(0) = S(0) - S(\infty) = -\int_0^\infty dS = \ldots = \int_0^\infty [E(\beta)_{\text{test}} - N]\, d\beta$$

- Peform $T = 0$ MC (orig. system) at temp. β (H_{test}), measure $\langle H_{\text{test}} \rangle_{\text{test}}$
- Adaptive choice of MC sweeps and 25 different values of β.



Hierarchical Clustering SAT

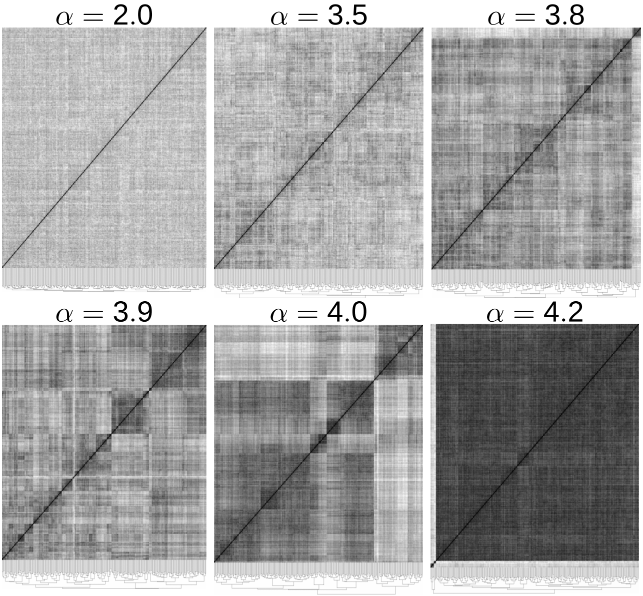
1 sample

($N = 256$)

small α:
1 cluster

medium α:
hierarchy
(# levels ?)

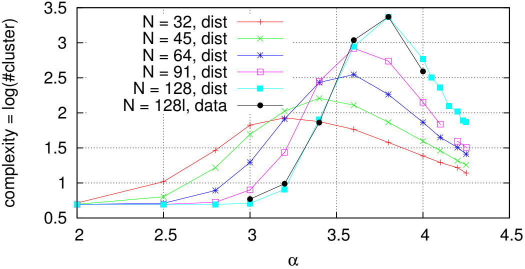
larger α:
landscape
simplifies



Results SAT

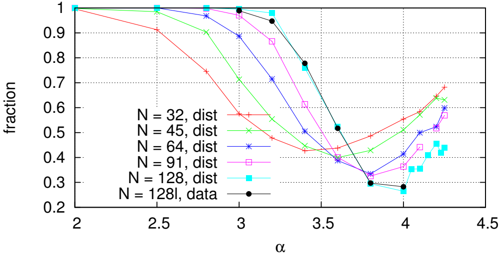
- Number of clusters (so far small N)

  (corr. to $\alpha_{d,+}$!?)



- Relative weight of largest cluster (so far small N)

  (corr. to α_d !?)

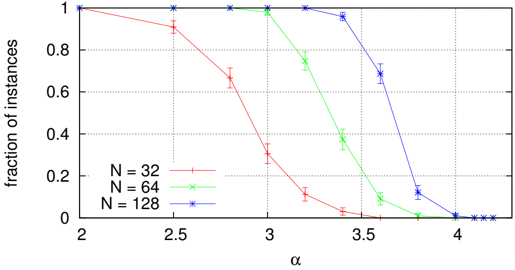


(data collaps using $\alpha' = (\alpha - \alpha_d)N^{1/\nu}$ for $\alpha_d = 4.1(1), \nu = 2.0(5)$)

Frozen Variables

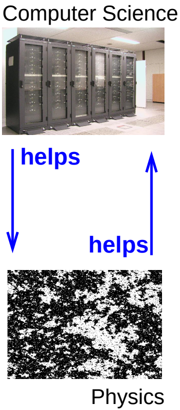
- Frozen variables:

  (makes it difficult for stochastic algorithms !?)



currently: looking directly at variables for $N \approx 10000$

Summary



- NP-hard combinatorial optimization problems:
  - vertex cover (VC)
  - satisfiability (SAT)
- Phase trans. on random ensembles
- Cluster analysis
  - Ballistic search/networking
  - Count number of clusters (neighbor based)
  - Hierarchical clustering ($\rightarrow$ dendogram, ordering)
- VC: complex structure (hard phase)
- SAT: several transitions

Finally ...

- Thank you for the attention !
- Post-doc position available
  - as soon as possible
  - for two years
  - computational physics: disordered systems, algorithms, optimization problems, computational complexity, large-deviation properties, biofinformatics, . . .
  - contact me: a.hartmann@uni-oldenburg.de