Analysis of the Stop-and-Restart procedure for solving random 3-SAT

A large deviation study of the Generalized Unit Clause search heuristic.

Simona Cocco *University Louis Pasteur, Strasbourg*

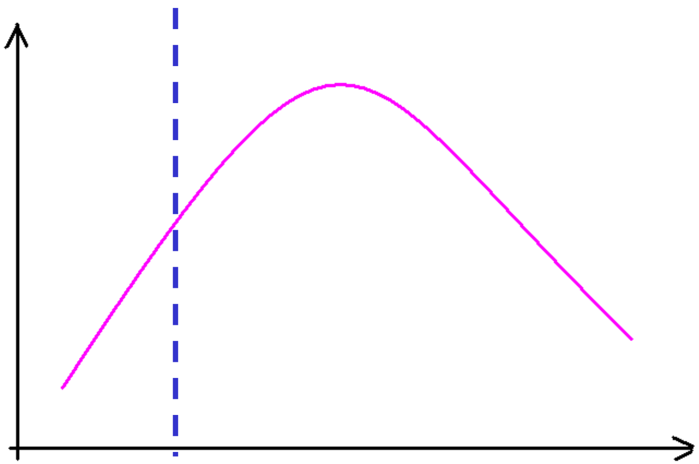
Rémi Monasson *Ecole Normale Supérieure, Paris*

http://www.lpt.ens.fr/~monasson

SAT 2002, Cincinnati

Introduction

Branch-and-bound algorithm fl non deterministic heuristic



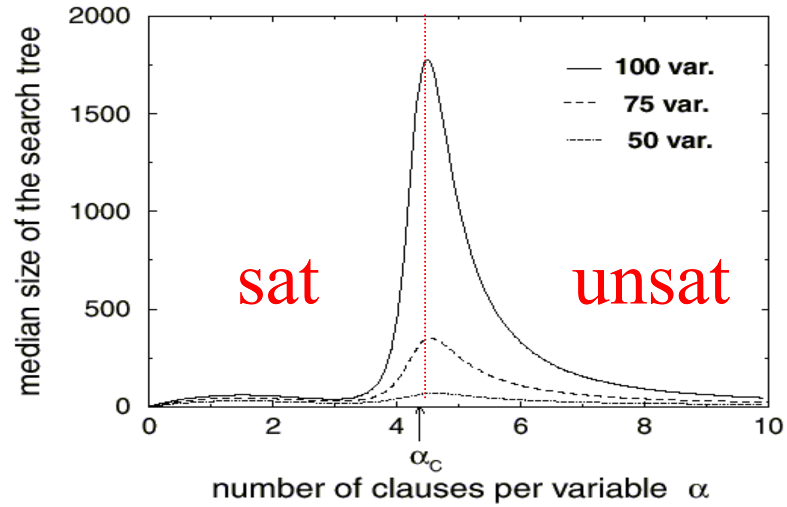
Idea: exploit the (left) tail of the distribution of resolution times

Stop & Restart ! (incomplete …)

Dubois et al. '93,
Gomes et al. '00

Random 3-SAT

$$\alpha = \frac{\text{nb. of clauses}}{\text{nb. of variables}} \rightarrow \boxed{\alpha_C \approx 4.3}$$



Davis-Putnam-Loveland-Logeman algorithm = heuristic + backtracking

Generalized unit-clause (GUC): pick literal in shortest clause

Mitchell, Selman, Levesque ’92; Crawford, Auton ’93; Gent, Walsh ’94; Chao, Franco ‘86, ’90; Chvatal, Szmeredi ’88; Achlioptas ’01 ….

Resolution trajectories

clauses with 3 var.

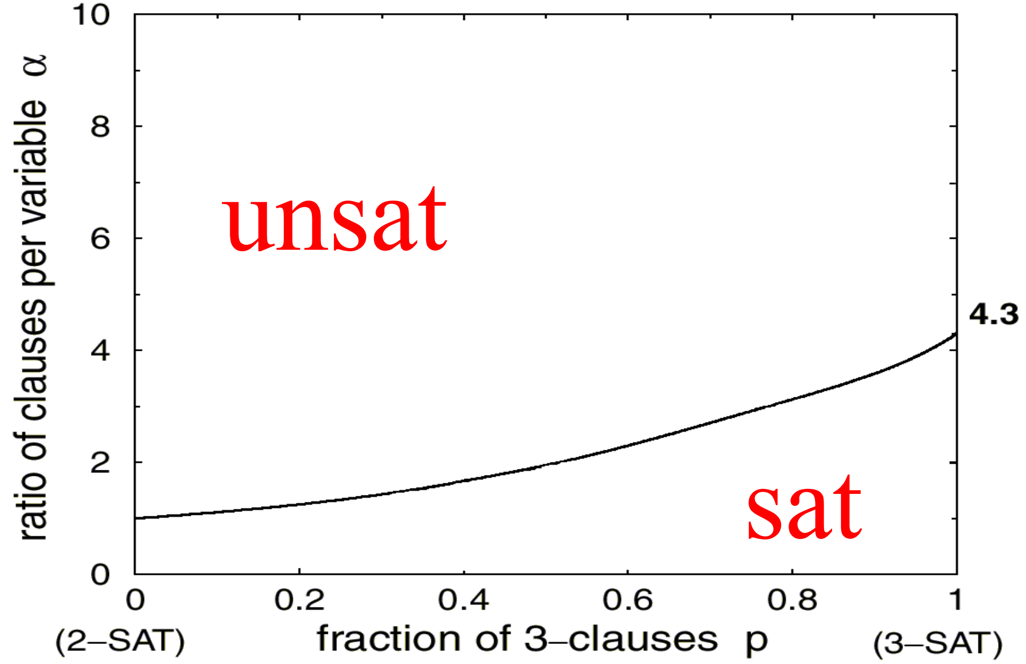
α

"dynamics" of the algorithm

clauses with 2 or 3 var.

α , p

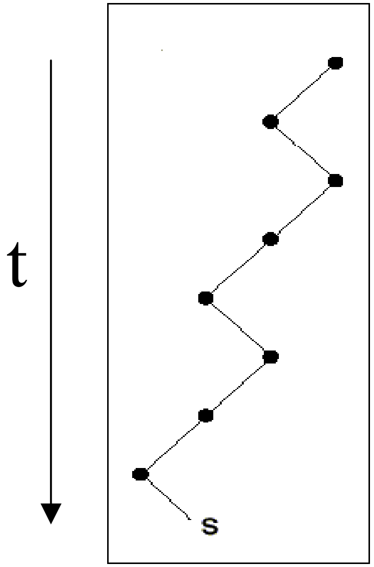
phase diagram of the 2+p-SAT model

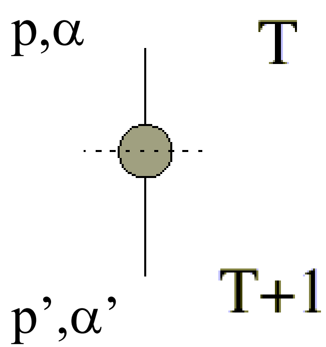


Monasson, Zecchina, Kirpatrick, Selman, Troyansky '99

Achlioptas, Kirousis, Kranakis, Krizanc '01

Satisfiable and easy instances $\alpha < 3.003$





$$p(0)=1\,,\ \alpha(0)=\alpha_0$$

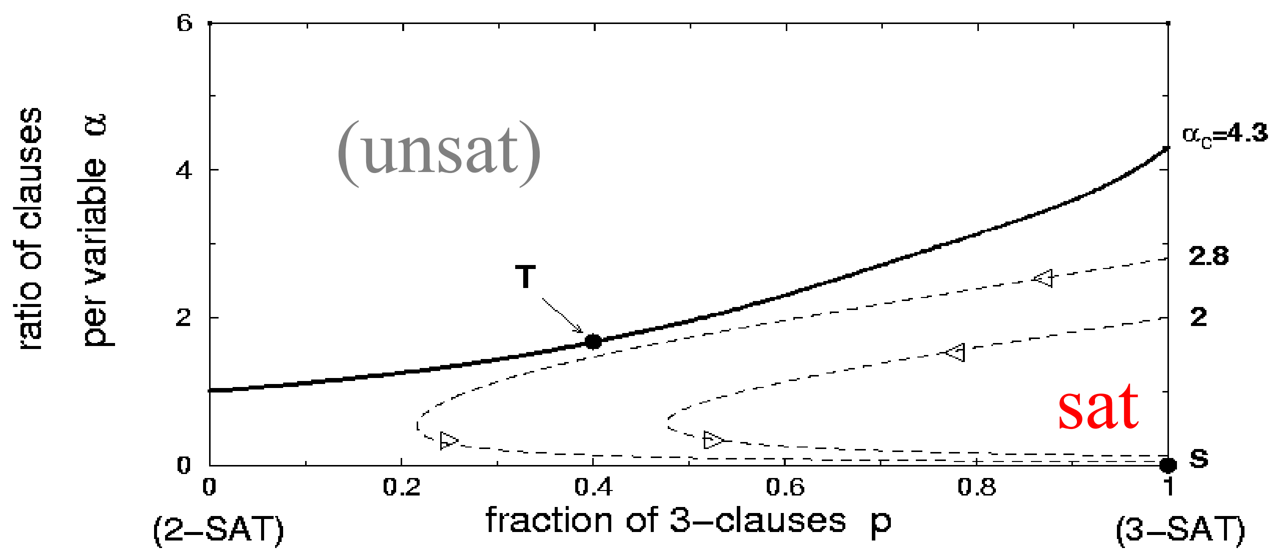
$$\frac{dp}{dt} = F_p(p,\alpha,t)$$

$$\frac{d\alpha}{dt} = F_\alpha(p,\alpha,t)$$

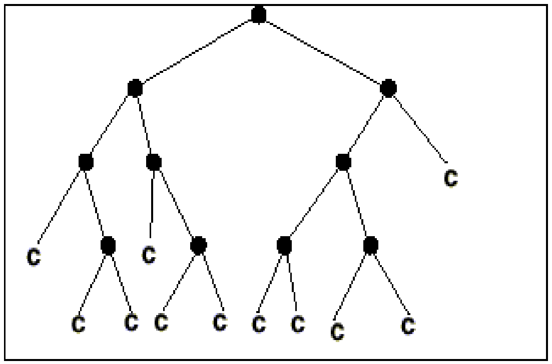
(ODE)

Chao, Franco '90

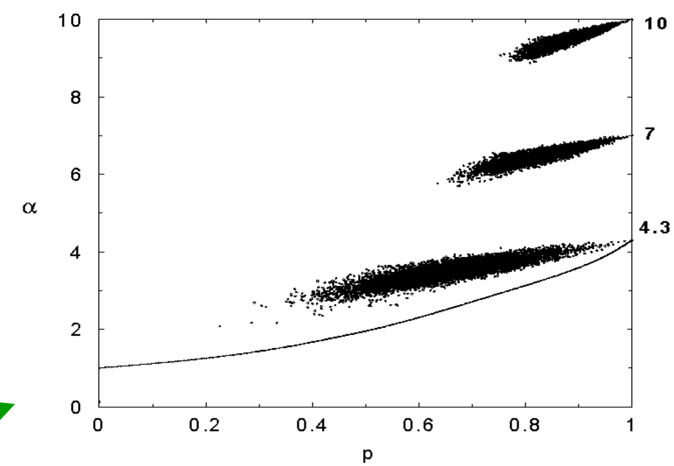
Frieze, Suen '96

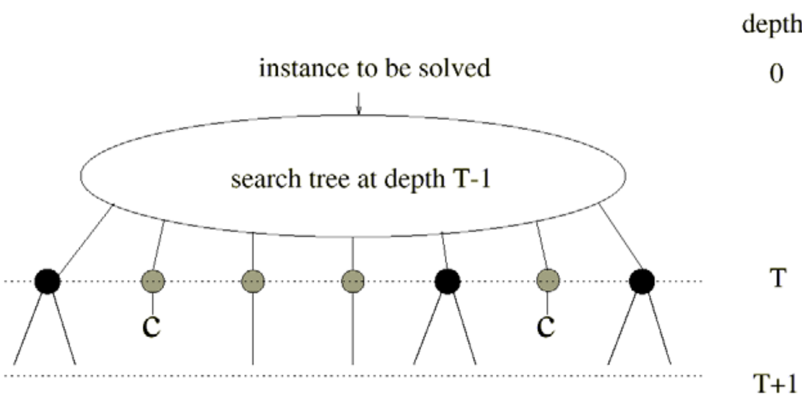


Unsatisfiable, hard instances $\alpha > 4.3$



DPLL induces a non Markovian evolution of the search tree

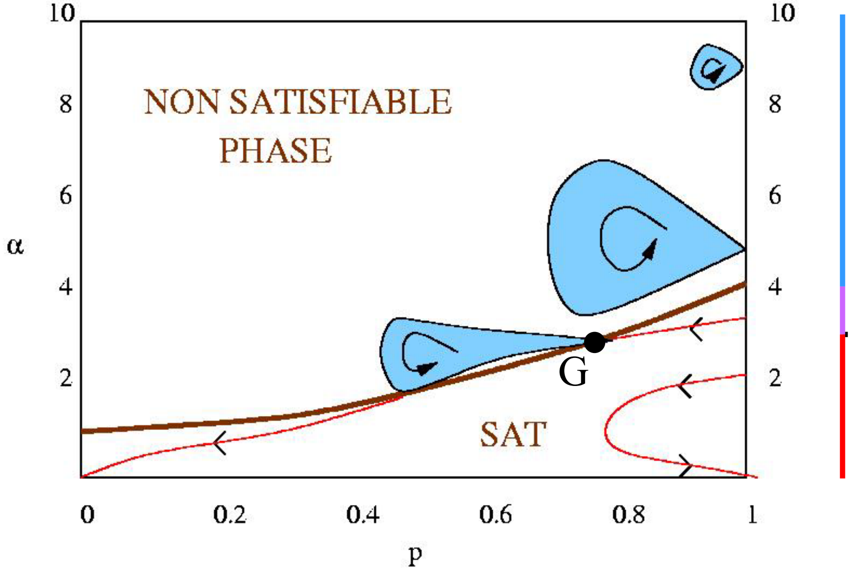


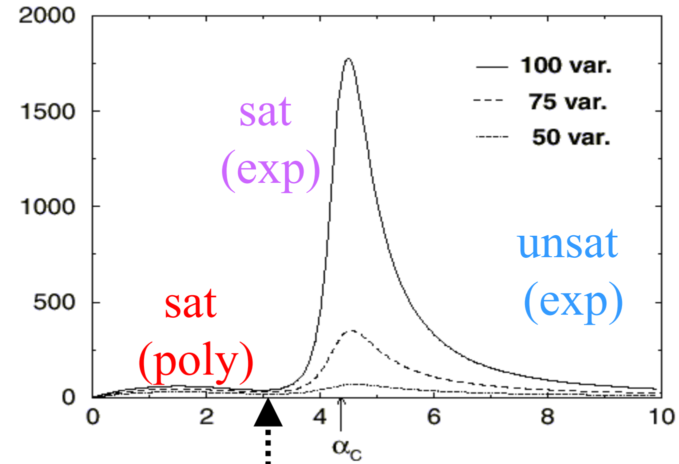


Imaginary, and parallel building up of the search tree

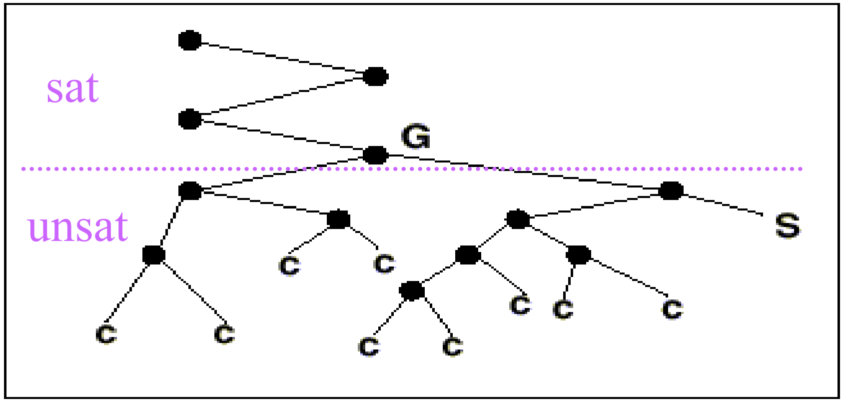
one branch: *$p(t)$, $\alpha(t)$* *ODE* → many branches: *$b\,(p, \alpha, t)$* *PDE*

The polynomial/exponentiel crossover





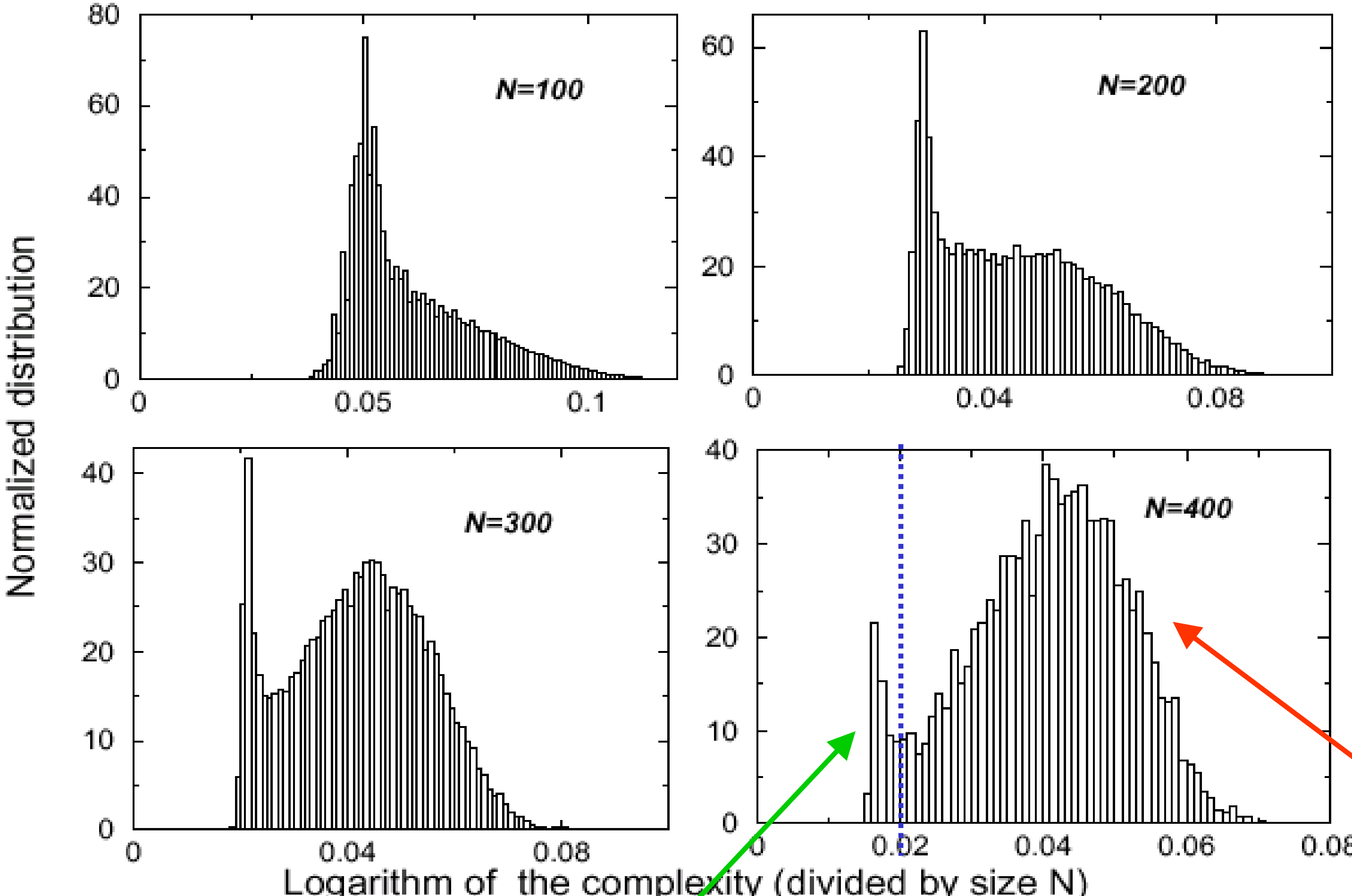
"dynamical" transition
(depends on the heuristic)



Satisfiable, hard instances 3.003< α < 4.3

Cocco, R.M. '01
Vardi et al. '00

Fluctuations of complexity for finite instance size



Histograms of solving times

$\alpha=3.5$

Exponential regime
Complexity $= 2^{0.035\,N}$

Linear regime
Very rare! frequence = $2^{-0.011\,N}$

Application to Stop & Restart resolution

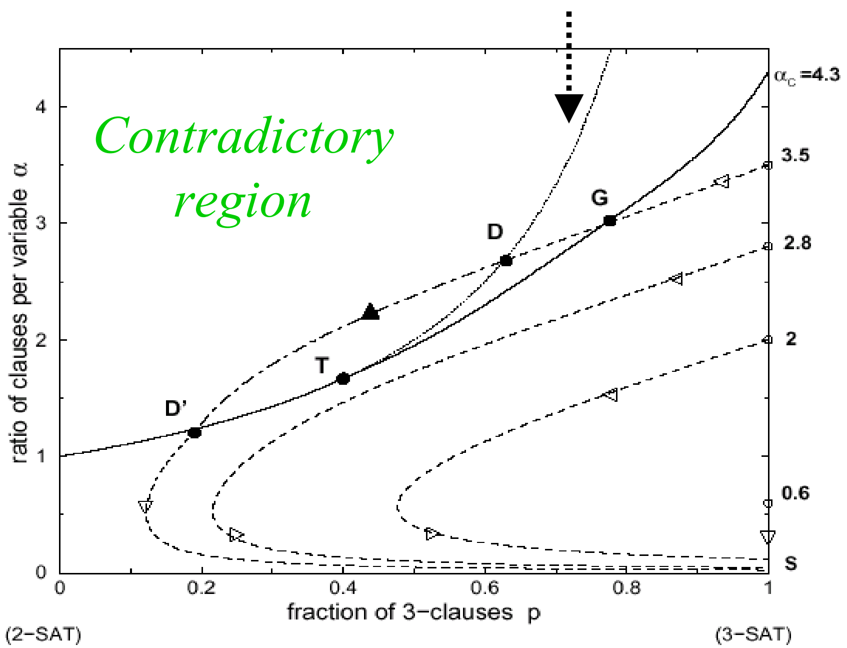
Resolution through systematic stop-and-restart of the search:

- stop algorithm after time N;
- restart until a solution is found.

Cocco, R.M. ‘02

Time of resolution : $2^{0.035\,N}$ → $2^{0.011\,N}$

Halt line for first branch = accumulation of unitary clauses



Easy resolution trajectories manage to survive in the contradictory region!

Analysis of the probability of survival *(I)*

Probability of survival of the first branch with clause populations C_1, C_2, C_3

$$B(C_1, C_2, C_3; T+1) = \sum_{C'_1, C'_2, C'_3 = 0}^{\infty} K\left(C_1, C_2, C_3, C'_1, C'_2, C'_3; T\right) B(C'_1, C'_2, C'_3; T)$$

Transition matrix

$$K(C_1, C_2, C_3; C'_1, C'_2, C'_3; T) = \binom{C'_3}{C'_3 - C_3} \left(\frac{3}{N-T}\right)^{C'_3 - C_3} \left(1 - \frac{3}{N-T}\right)^{C_3} \sum_{w_2=0}^{C'_3 - C_3} \left(\frac{1}{2}\right)^{C'_3 - C_3} \binom{C'_3 - C_3}{w_2}$$

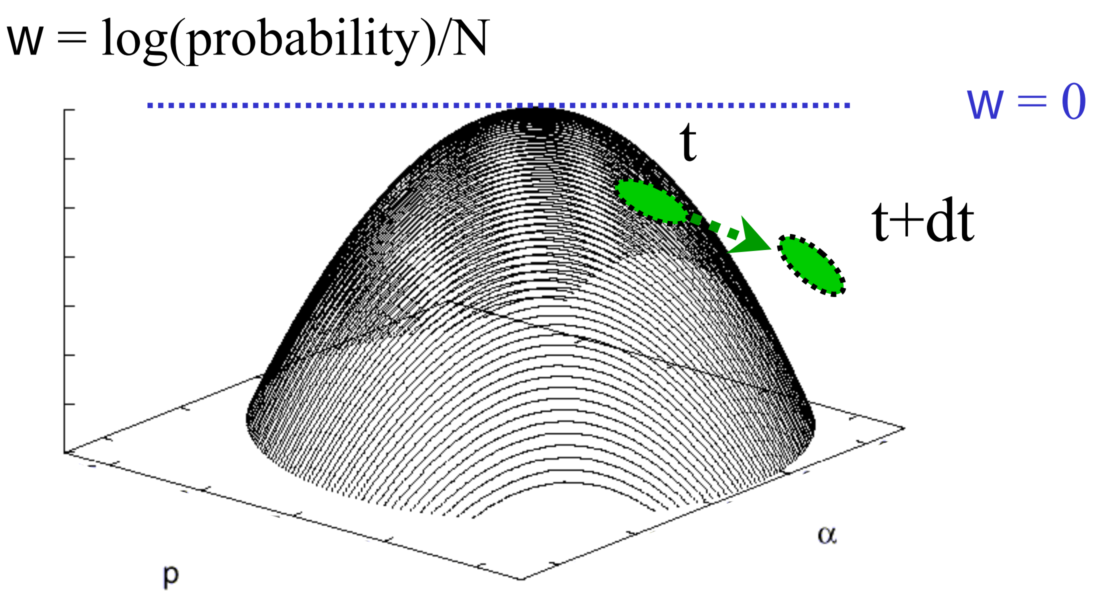
$$\times \Bigg\{ (1 - \delta_{C'_1}) \sum_{z_2=0}^{C'_2} \binom{C'_2}{z_2} \left(\frac{2}{N-T}\right)^{z_2} \left(1 - \frac{2}{N-T}\right)^{C'_2 - z_2} \sum_{w_1=0}^{z_2} \left(\frac{1}{2}\right)^{z_2} \binom{z_2}{w_1} \delta_{C_2 - C'_2 - w_2 + z_2} \, \delta_{C_1 - C'_1 - w_1 + 1} + \delta_{C'_1}$$

$$\times \sum_{z_2=0}^{C'_2 - 1} \binom{C'_2 - 1}{z_2} \left(\frac{2}{N-T}\right)^{z_2} \left(1 - \frac{2}{N-T}\right)^{C'_2 - 1 - z_2} \sum_{w_1=0}^{z_2} \left(\frac{1}{2}\right)^{z_2} \binom{z_2}{w_1} \delta_{C_2 - C'_2 - w_2 + z_2 + 1} \, [\delta_{C_1 - w_1} \; {}^{*}(1 - C_1/2/(N-T))^{C1-1} \Bigg\},$$

- $B(C_1, C_2, C_3; T) \sim \exp[-N \; w(c_1, c_2, c_3; t)]$ where $c_i = C_i/N$, $t = T/N$
- two cases: $C_1 = O(1)$ (safe regime), $C_1 = O(N)$ (dangerous regime).

Analysis of the probability of survival *(II)*

- *Safe regime:*
  $c_1 = 0$



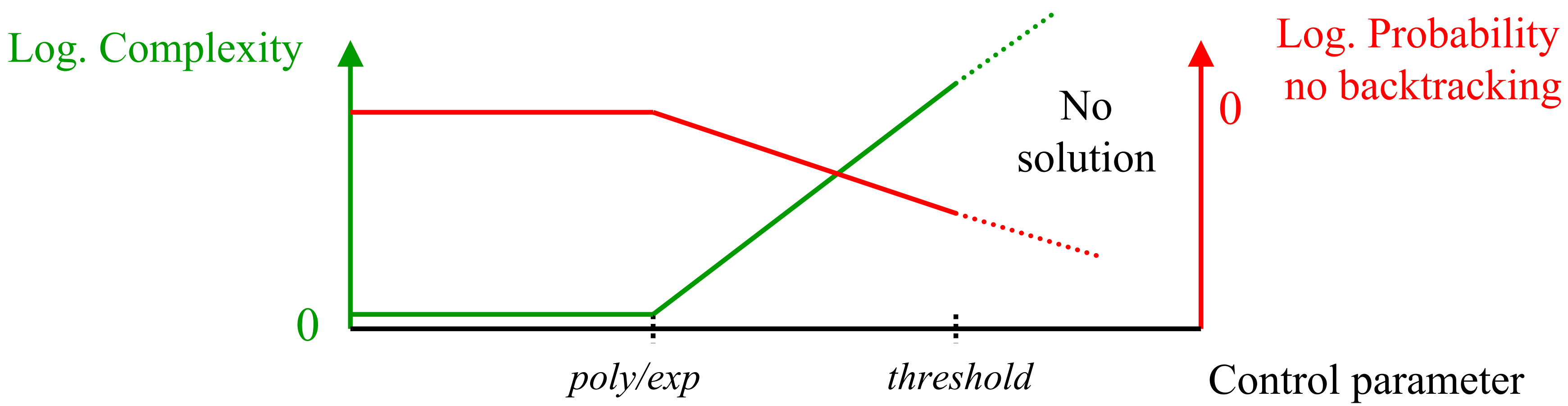
$$\frac{\partial \omega}{\partial t} = \mathcal{H}\left[p, \alpha, \frac{\partial \omega}{\partial p}, \frac{\partial \omega}{\partial \alpha}, t \right] \qquad \text{(PDE)}$$

- *Dangerous regime:* $c_1 = O(1)$, $w < 0$

Conclusions

- Statistical physics concepts and techniques can be useful to reach a non rigorous but quantitative understanding of algorithms
  (phase diagram, dynamical trajectories, growth processes, ...)

- Stop-&-Restart procedure: when do we stop?
  from the present study: cut-off time = size of the instance



- Further developments:
  analysis of trees generated by more complex heuristics, ...