Analyse en moyenne de problèmes NP-complets : pourquoi tant d'algorithmes ont une complexité $n^{c \log n}$?

Cyril Banderier, Journées combinatoires de Bordeaux

(en collaboration avec Hsien-Kuei Hwang, Vlady Ravelomanana, Vytas Zacharovas)

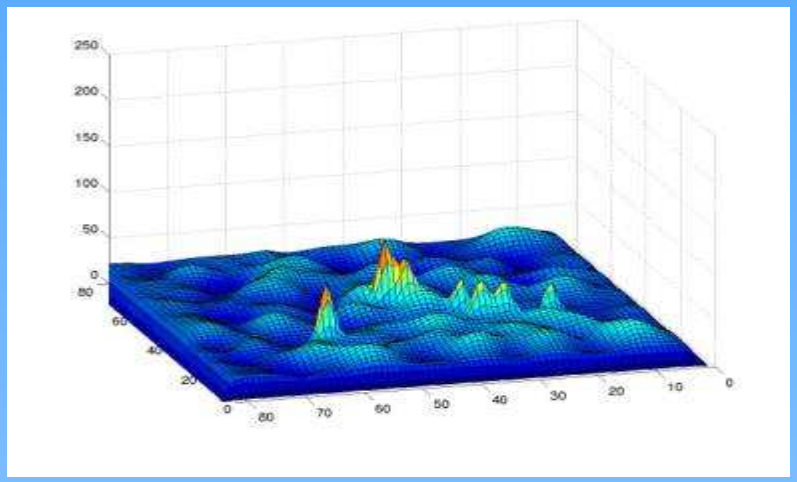
13 février 2013




Plan of the talk

- **Complexity landscapes**
- **A constraint satisfaction problem (CSP): Maximum independent set (MIS)**
- **Its average complexity (asymptotics)**
- **Ubiquity of $n^{\ln n}$**
- **Mathematics behind**

Complexity landscape of algorithms



Worst-case complexity: $\mathrm{Max}_x\, T(x)$

Average-case complexity: $\mathrm{E}_x[T(x)]$

Smoothed complexity: $\mathrm{Max}_x \mathrm{E}_\epsilon[T(x+\epsilon)]$

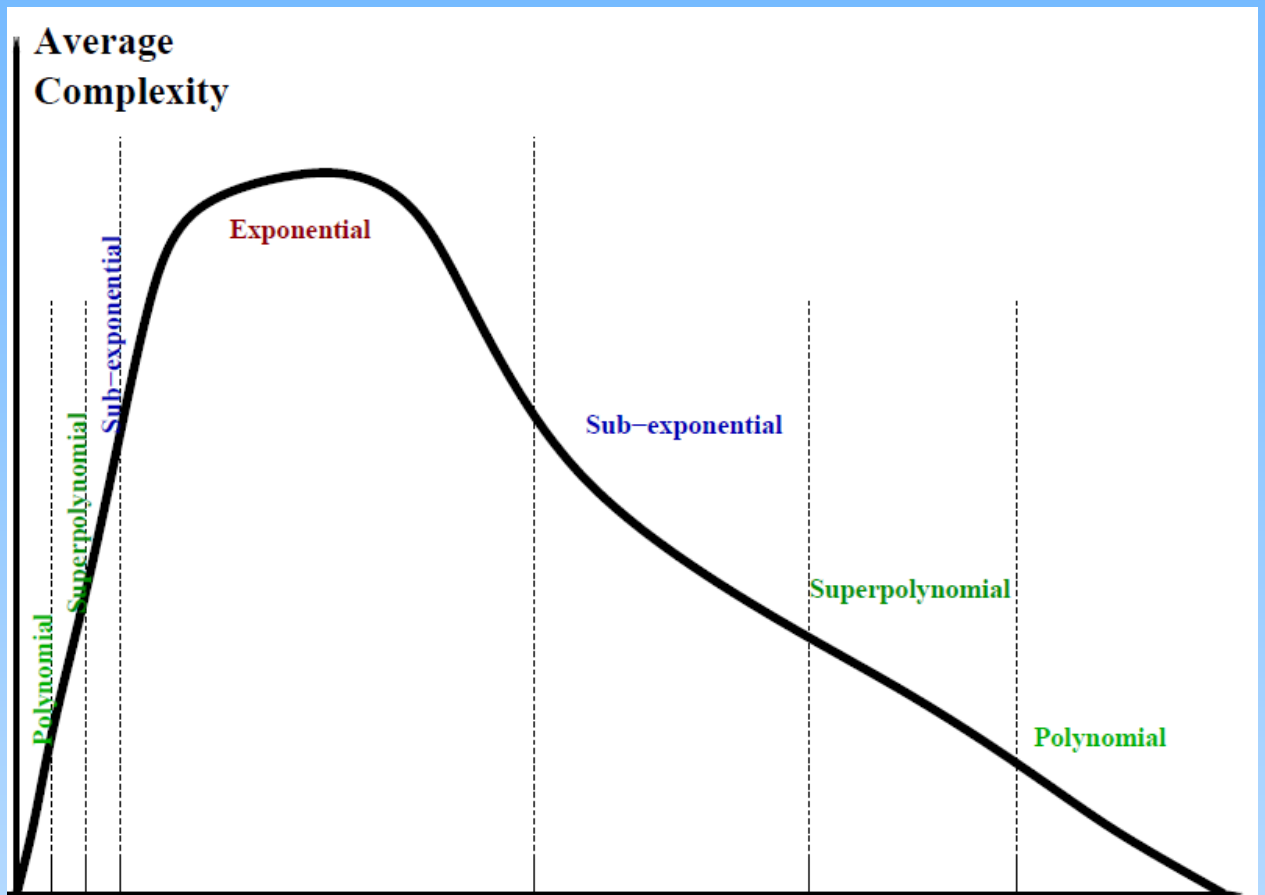
$\longrightarrow$ **what are the difficult regions?**

Smoothed Analysis of Algorithms (∼20 articles so far):

- **Why The Simplex Algorithm Usually Takes Polynomial Time [Spielman & Teng 01]**
- **Smoothed analysis of three combinatorial problems [Banderier & Beier & Mehlhorn 03]**

Complexity landscape of graph algorithms... (with respect to the density)

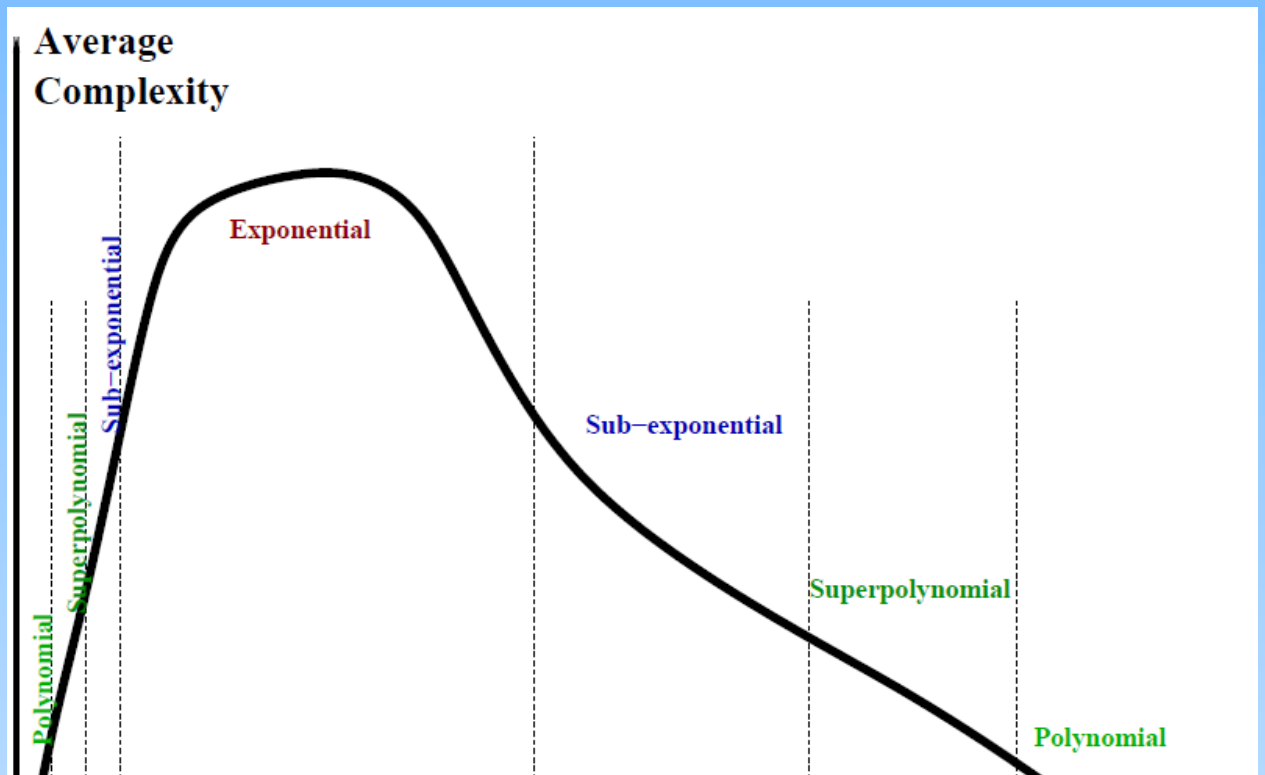
For discrete combinatorial structures: no unique notion of perturbation/vicinity/neighboorhoud".
$\rightarrow$ A natural measure of "similarity" for graphs: the density $\frac{\#\text{ edges}}{\#\text{ vertices}}$.



Complexity landscapes and $P \neq NP$

If one could prove a superpolynomial average complexity for a problem (and not only for an algorithm), this would prove $P \neq NP$.

We prove such a superpolynomial average complexity for a infinite class of algorithms.



Introduction to Constraint Satisfaction Problems

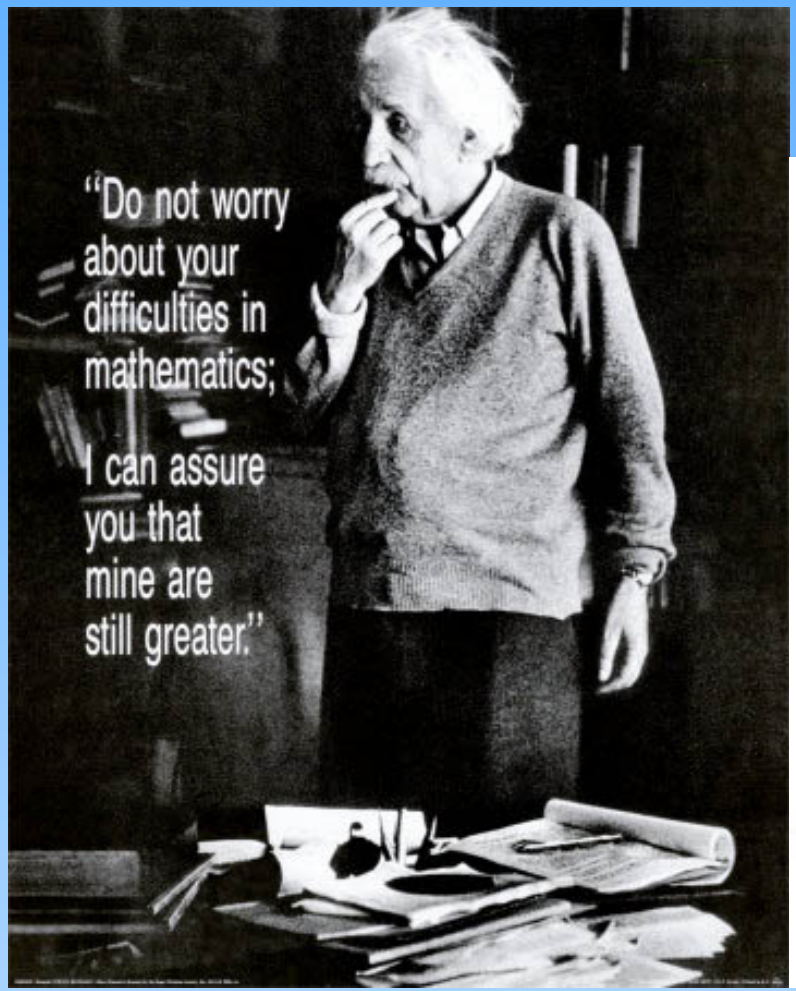
Constraint Satisfaction Problems

The class of problems known as Max–CSP plays a key rôle in Computer Science.

CSP

- n **variables** $x_1, \cdots, x_n$ **belonging to finite domains.**
- **A set of constraints (clauses) over these variables: a constraint is a relation between the variables defining** *authorized* **combinations of them.**
- **Decision problem: does it exist an assignment of the variables (aka, a solution) satisfying all the constraints?**
- **Optimization problem: maximize the number of satisfied constraints.**

Constraint Satisfaction Problem: Einstein Problem



Constraint Satisfaction Problem: Einstein Problem

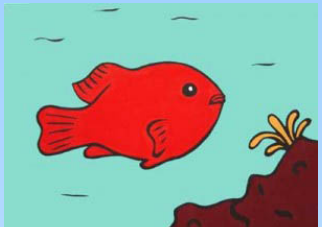
The author of this problem is Albert Einstein who said that 98% of the people in the world couldn't solve it.

Facts:

1. There are 5 houses (along the street) in 5 different colors: blue, green, red, white and yellow.
2. In each house lives a person of a different nationality: Brit, Dane, German, Norwegian and Swede.
3. These 5 owners drink a certain beverage: beer, coffee, milk, tea and water, smoke a certain brand of cigar: Blue Master, Dunhill, Pall Mall, Prince and blend, and keep a certain pet: cat, bird, dog, fish and horse.
4. No owners have the same pet, smoke the same brand of cigar, or drink the same beverage.

Hints:

1. The Brit lives in a red house.
2. The Swede keeps dogs as pets.
3. The Dane drinks tea.
4. The green house is on the left of the white house (next to it).
5. The green house owner drinks coffee.
6. The person who smokes Pall Mall rears birds.
7. The owner of the yellow house smokes Dunhill.
8. The man living in the house right in the center drinks milk.
9. The Norwegian lives in the first house.
10. The man who smokes blend lives next to the one who keeps cats.
11. The man who keeps horses lives next to the man who smokes Dunhill.
12. The owner who smokes Blue Master drinks beer.
13. The German smokes Prince.
14. The Norwegian lives next to the blue house.
15. The man who smokes blend has a neighbor who drinks water.



Who keeps the fish?

Constraint Satisfaction Problem: a less sexy example

Domains: $(x, y) \in \{0, 1\}^2, (z, t) \in \{0, 1, 2, 3\}^2$

Constraints:

$$\begin{cases} (x, y) & \in & \{(1, 0), (0, 1)\} \\ x & \neq & z \\ y + z & = & 0 \mod 2 \\ t & \geq & y \end{cases}$$

Definition: Max-2-CSP

- **Each constraint concerns at most 2 variables.**

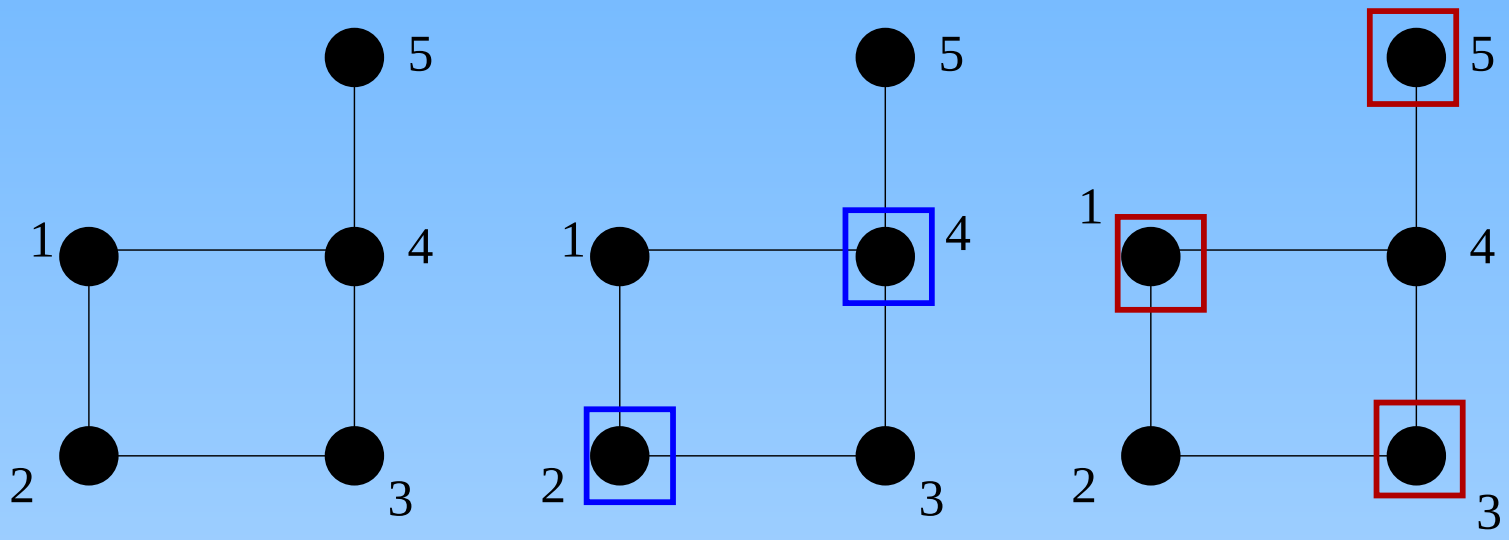
Examples of Max-2-CSP problems

Max-2-CSP is very general paradigm:

- **Maximum bipartite subgraph (or Max-2-COL)**
- **Max-CUT**
- **Max-2-SAT, Max-2-XORSAT**
- **graph colorability...**
- **Maximum Independent Set (MIS).**

IS (independent set) and MIS (maximum independent set)

IS = set of nodes not linked together.
MIS = set of nodes not linked together and one cannot get larger.



{2, 4} is an IS

{1,3,5} is an IS
of MAXIMUM size

N.B.: MIS not always unique.

Why the MIS problem?

Stable maximum (Maximum Independent Set or MIS).

Motivations

1. **Numerous applications** (coloration, finances, codes, distributed systems, ...)
2. **The typical NP-complete problem (cf. Karp's original list [KARP 72]).**
3. **NP-HARD to approximate Håstad 99.**
4. **Many works including those of [TARJAN 77, CHVÀTAL 77, PITTEL 82, BALAS, YU 86, KARMARKAR, RESENDE 89, ...,] .**

Finding a MIS: some remarkable works

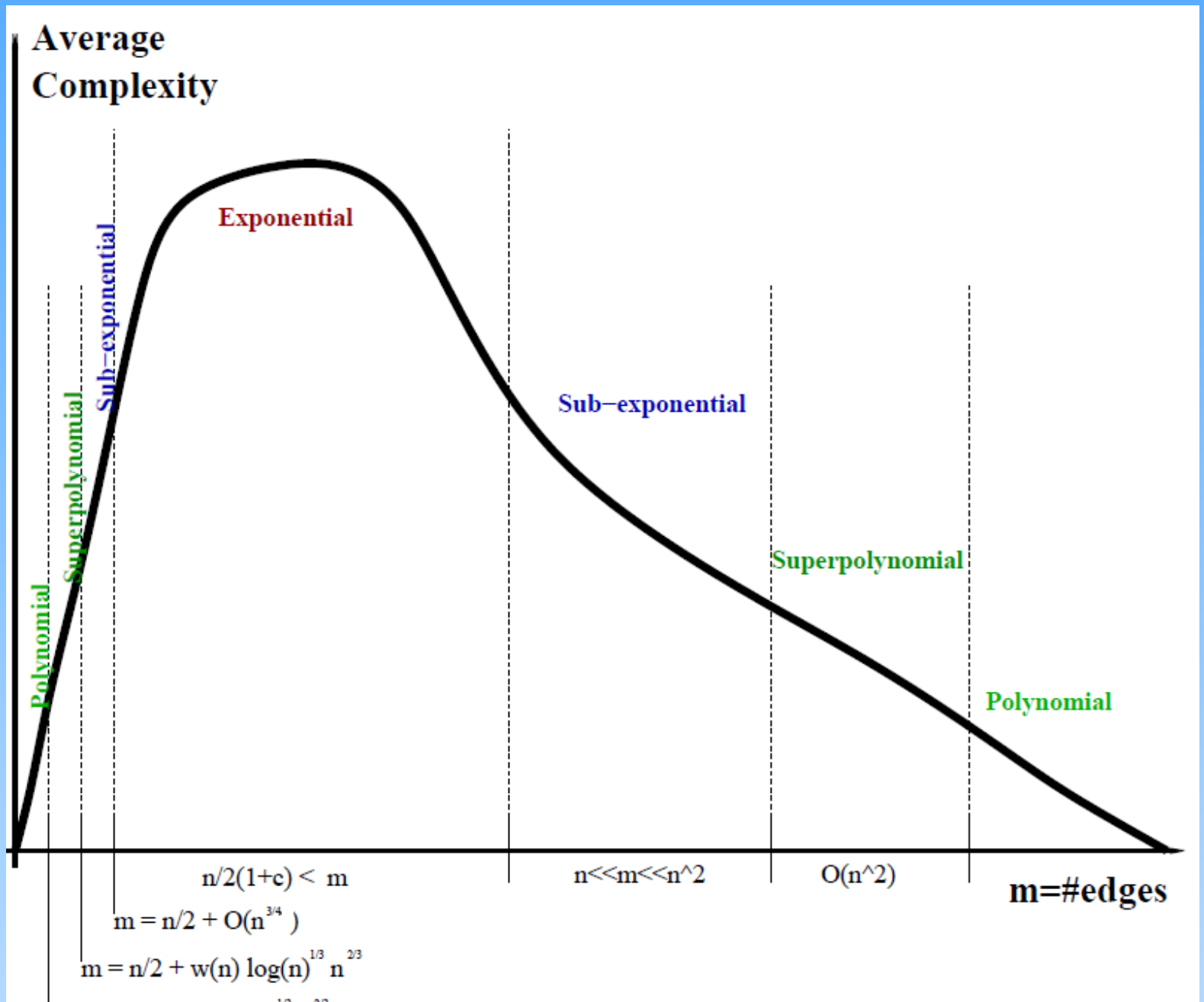
1. **[TARJAN, TROJANOWSKI 77] A seminal paper for worst-case analysis: introduce the idea of reductions.**

2. **[CHVÀTAL 77] Complexity theory (Length of recursive proofs) for finding a MIS. W. h. p., lengths are of order $e^{\Theta(\log n^2)}$.**

3. **[PITTEL 82] W. h. p. when the graph is dense, $e^{\Theta(\log n^2)}$ iterations are required.**

OUR WORK: More precise average-case analysis using various analytical approaches: generating functions, Laplace transforms, ..., for a full class of algorithms, for any density.

OVERVIEW OF OUR RESULTS



Main ideas of the proof

1. **WEST SIDE.**
   We analyze a very general algorithm working on all MAX-2-CSP. We use reductions (transformations of the underlying graph) mainly on vertices of degree < 3. We "`branch`" only on vertices of degree ≥ 3.

2. **EAST SIDE.**
   We analyze an exhaustive algorithm.

Remark.

In both cases, the algorithms are exact and ALWAYS RETURN a MIS. As input, the underlying graphs of the constraints are random $\mathbb{G}(n, m)$ and $\mathbb{G}(n, p)$. Our results quantify the average number of global iterations of the algorithms.

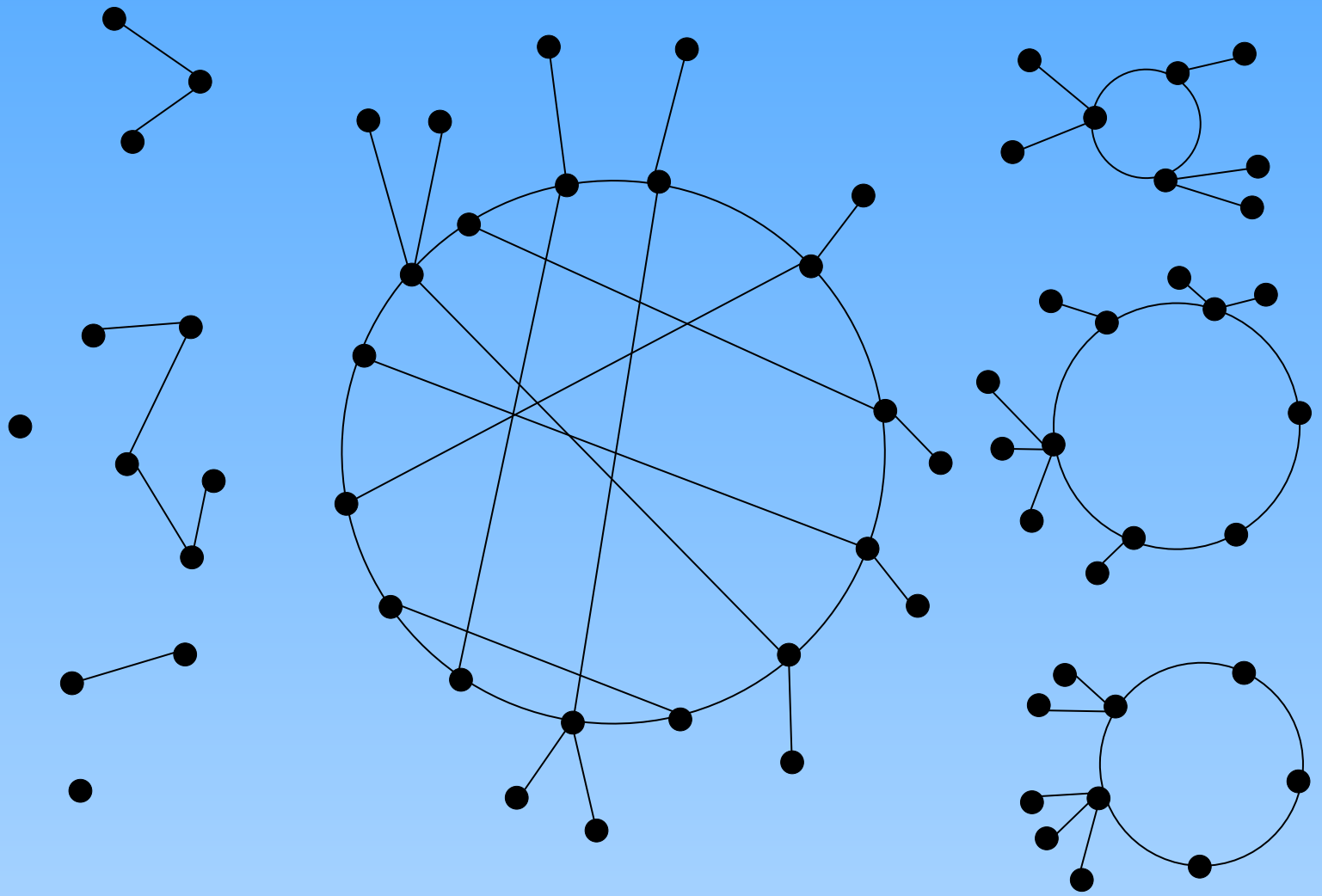
West Side Story

WEST SIDE: algorithm for MAX-2-CSP, main ideas and analysis

- **Reduce all vertices of degree 1 and 2. (as in [TARJAN 77]).**
- **These reductions cost polynomial time.**
- **Branch on vertices of degree ≥ 3.**
- **On graphs of excess r, the cost is $< \sqrt{2}^r$ ([SCOTT, SORKIN 03]**
- **The average cost for graphs of $\mathbb{G}(n, m)$ is therefore upper bounded by**

$$\sum_{r=0}^{m} 2^{r/2} \underbrace{\texttt{Proba}\,[\,\textbf{graph } n \textbf{ vertices }, m \textbf{ edges has excess } r]}_{p_r(n,m)}$$

WEST SIDE: graph before the reductions

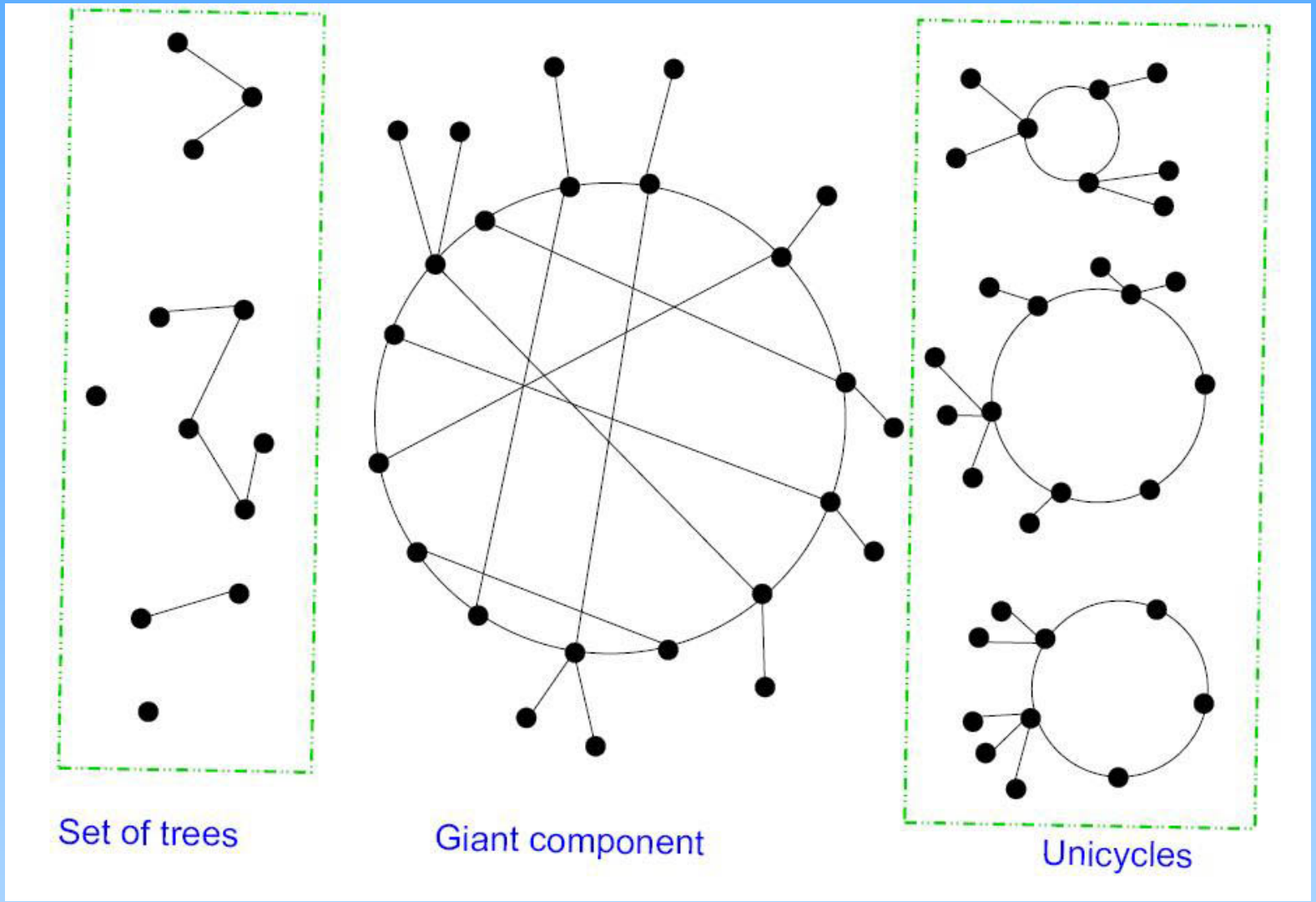


Set of trees

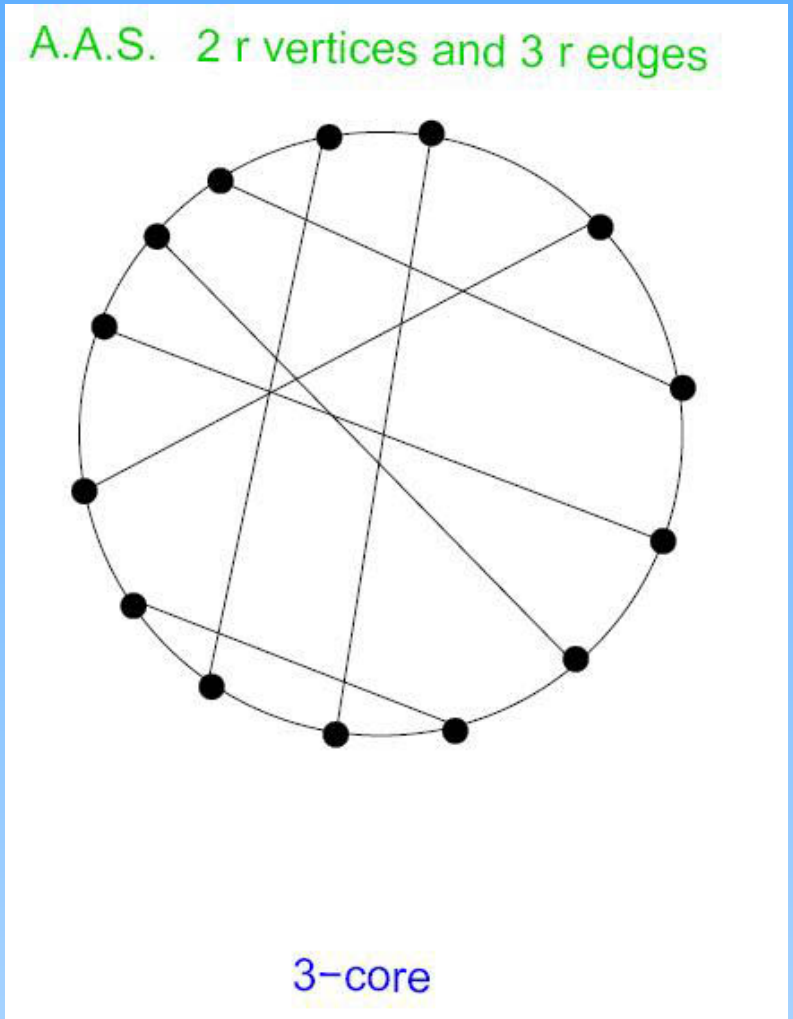
Giant component

Unicycles

WEST SIDE: graph before the reductions



WEST SIDE: graph after reductions



WEST SIDE: computing $p_r(n,m)$

Theorem. **The proba. that a graph with n vertices and $m = \frac{n}{2} + \mu n^{2/3}$ edges has a connected component of excess $r = \frac{16}{3}\mu^3 + O(\mu^{3/2})$ is**

$$\mathbf{p_r(n,m)} \sim \left(\frac{\mathbf{3}}{\mathbf{160}\pi\mu^{\mathbf{3}}}\right)^{\mathbf{1/2}} \exp\left(-\frac{\mathbf{3}}{\mathbf{160}}\frac{\left(\mathbf{r}-\frac{\mathbf{16}}{\mathbf{3}}\mu^{\mathbf{3}}\right)^{\mathbf{2}}}{\mu^{\mathbf{3}}}\right)$$

(with uniform error terms as $r,\ \mu,\ n \to \infty$ et $|r - \frac{16}{3}\mu^3| \le O(\mu^{3/2})$, $r \le O(n^{1/4})$ and $\mu \le O(n^{1/12})$).

Proof. Use GF (here $T \equiv T(z) = ze^{T(z)}$)

$$p_r(n,m) \sim \frac{n!}{\binom{\binom{n}{2}}{m}} [z^n] \underbrace{\frac{(T(z) - T(z)^2/2)^{n-m}}{(n-m)!}}_{\text{trees}} \underbrace{\frac{e^{-T(z)/2 - T(z)^2/4}}{(1-T(z))^{1/2}}}_{\text{unicycles}} \underbrace{W_r(z)}_{\text{giant cor}}$$

where W_r is Wright's GF for graphs of excess r (+saddle-point method [KNUTH-JANSON-ŁUCZAK-PITTEL 93, FLAJOLET-SEDGEWICK 09]).

WEST SIDE: results

Recall that the desired average cost of the reduction algorithm is

$$\sum_{r=0}^{m} 2^{r/2} \underbrace{\texttt{Proba}\,[\text{ graph } n \text{ vertices }, m \text{ edges has excess } r]}_{p_r(n,m)} .$$

Theorem. **Let $m = \frac{n}{2} + \mu n^{2/3}$ with $\mu = o(n^{1/3})$. The average cost of ALGO on $\mathbb{G}(n,m)$ is at most**

$$\exp\left(\frac{2\log 2}{3}\,(5\log 2 + 4)\,\mu^3\right)\Big(1 + o(1)\Big)\,.$$

If $m = \frac{n}{2}(1+\varepsilon)$ with $n^{-1/4} \le \varepsilon < \delta$, the average cost of ALGO is at most

$$\exp\left(\frac{5(\log 2)^2}{12}\varepsilon^3 n + \frac{\log 2(\varepsilon^2 - \sigma^2)}{4(1+\varepsilon)} n\right)\Big(1 + o(1)\Big)$$

with σ defined by $(1+\varepsilon)e^{-\varepsilon} = (1-\sigma)e^{\sigma}$.

Our theorem holds for all MAX-2-CSP problems and not only for ALGO–MIS!

East Side Story: Exhaustive Algorithms

EAST SIDE: an exhaustive algorithm

Procedure **MAXIS**$(G = (V, E) :$ `labeled graph`$)$
Pick the least vertex v **in** V **;**

(⋆ EITHER v **is not in the MIS** $\longrightarrow$ **remove it** $\star$**)**
$V_1 := V \setminus \{v\}$**;**
$E_1 := E \setminus \{$ **incident edges of** $v\}$ **;**
maxis$_1$ $:=$ **Card**(**MAXIS**$(G_1))$ **;**

(⋆ OR v **is in a MIS** $\longrightarrow$ **remove** v **and all its neighbors** $\star$**)**
$V_2 := V \setminus \Gamma(v)$**;** **(⋆** $\Gamma(v) = \{v\} \bigcup$ **its neighbors** $\star$**)**
$E_2 := E \setminus \{$ **incident edges of the vertices in** $\Gamma(v)\}$ **;**
maxis$_2$ $:= 1 +$ **Card**(**MAXIS**$(G_2))$ **;**

RETURN: $\max$ (**maxis**$_1$, **maxis**$_2$)

EAST SIDE: analysis of the exhaustive algorithm (1/2)

Let $M_n =$ be the mean number of iterations of the algorithm whenever the input are graphs from $\mathbb{G}(n,p)$. We have

$$\mathbf{M_n = M_{n-1} + \sum_{i=0}^{n-1} \binom{n-1}{i} p^i (1-p)^{n-1-i} M_{n-1-i}}$$

Theorem. **As $p \to 0$,**

$$\mathbf{M_n = \frac{1 - \frac{W(np-p)}{np-p}}{1 + W(np-p)} \left(1 - \frac{W(np-p)}{np-p}\right)^{-n} (1 + O(p)),}$$

where W is the Lambert function.

Example 1: $p = \frac{\log n}{n}$. Roughly $\exp\left(n \frac{\log\log n}{\log n}\right)$ iterations.

Example 2: $p = \frac{1}{n^{1/2}}$. Roughly $\exp\left(\frac{1}{2} n^{1/2} \log n\right)$ iterations.

Sub-exponential!

EAST SIDE: analysis of the exhaustive algorithm (2/2)

Theorem. **If** $0 < p < 1$ **(fixed),**

$$\mathbf{M_n} \sim \frac{\mathbf{e}^{(\rho/2)(\log(1/r))^2}\mathbf{G}(\rho\log(1/r))}{\mathbf{r}^{\rho+1/2}\sqrt{2\pi\rho\log(1/r)}},$$

where $\rho = 1/\log(1/(1-p))$, $r = \frac{W(n/\rho)}{n/\rho}$ **and** G **is continuous and 1-periodic function:**

$$G(u) = q^{(\{u\}^2+\{u\})/2} \sum_{-\infty<j<\infty} \frac{q^{j(j+1)/2}}{1+q^j q^{-\{u\}}}\, q^{-(j+1)\{u\}}.$$

Roughly: $n^{O(\log n)}$**. Superpolynomial!**

Discussions

Robustness of our classifications

- **These very simple algorithms are amongst the fastest ones.**
- **Our classifications are robust:**
  1. **Most of existing algorithms are "Tarjan-Chvàtal-like":**

$$T(n) \leq T(n-6) + T(n-4) \longrightarrow O((1+\varepsilon)^n)\,.$$

  2. **Even if one day a genious finds a fast algorithm leading to a recurrence of the type**

$$T(n) = T(n-A) + T(\lceil cn \rceil), \qquad A \text{ fixed and } c < 1$$

or

$$T(n) = T(n-A) + \sum_{i=0}^{n-A-1} \binom{n-A-1}{i} p^i (1-p)^{n-A-1} T(i)\,,$$

all of them are expected to lead to a superpolynomial complexity!!!!

Open problems

1. **A WELL-KNOWN and LONGSTANDING PROBLEM (cf. [KARP] and [FRIEZE] 70++):**

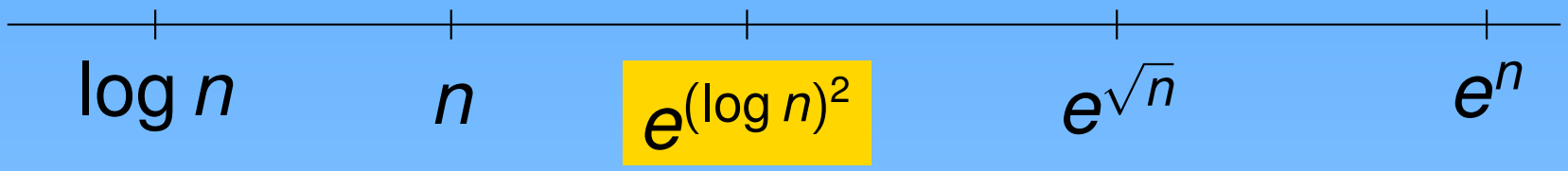
   Is there a POLYNOMIAL TIME algorithm that finds a MIS in $\mathbb{G}(n, p = 1/2)$?

2. **OUR WORK suggests an intermediate CHALLENGE:**

   Is there a $n^{o(\log n)} = n^{\text{SMALL-OH}(\log n)}$ algorithm which finds a MIS in $\mathbb{G}(n, p = 1/2)$?

   Recall that the Metropolis algorithm [JERRUM] "works" in $n^{O(\log n)} = n^{\text{BIG-OH}(\log n)}$.

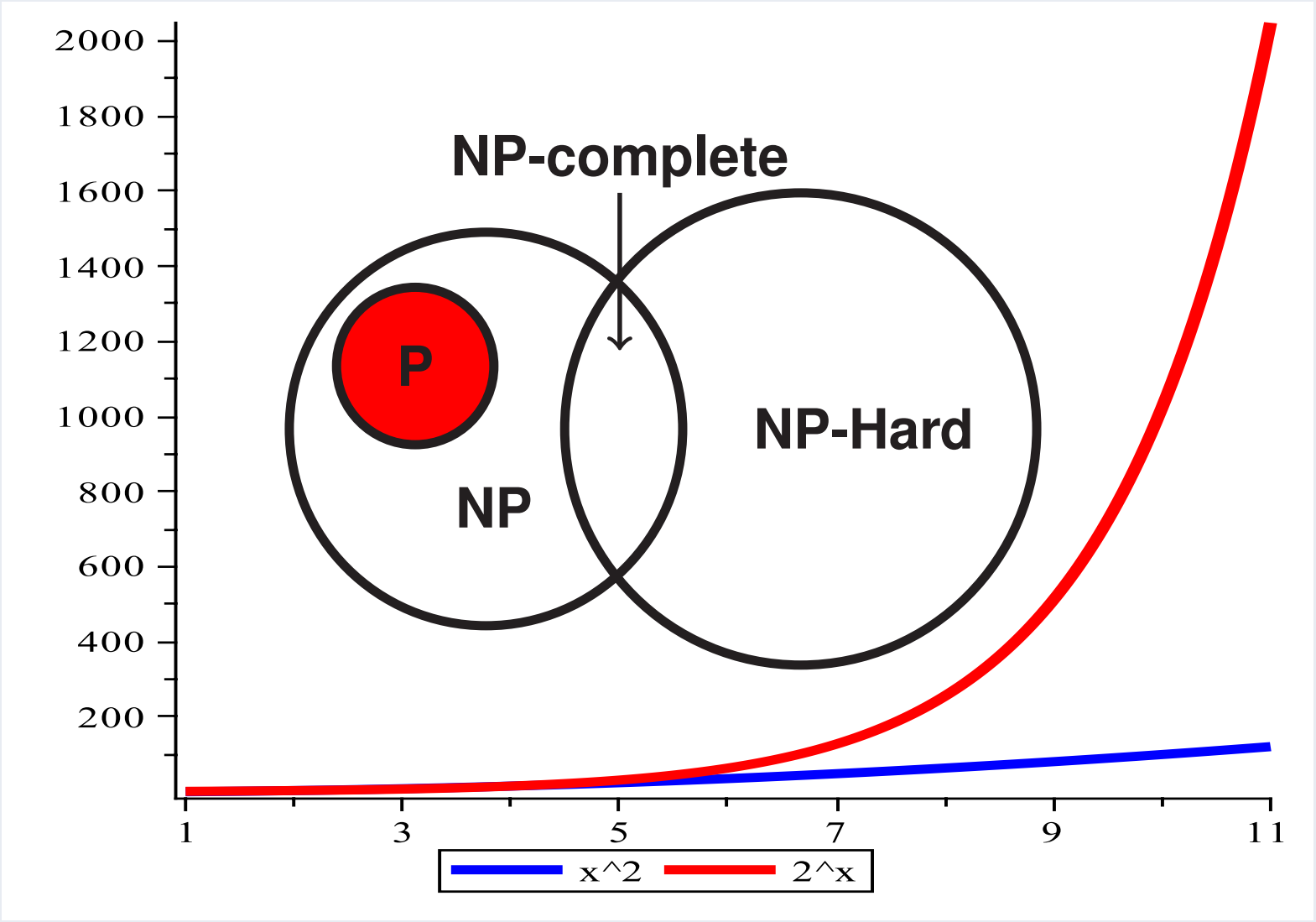
$$n^{c \log n} = e^{c(\log n)^2}$$



Ubiquity of the scale $n^{\ln(n)}$ in several worlds

POLYNOMIAL VS EXPONENTIAL

n^{α} vs α^{n}



LOG-NORMAL DISTRIBUTION

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma x} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right)$$

WIKIPÉDIA
L'encyclopédie libre

Accueil
Portails th é matiques
Index alphab é tique
Article au hasard
Contacter Wikip é dia

Contribuer
Premiers pas
Aide
Communaut é
Modifications r é centes
Faire un don

Imprimer / exporter

Boîte à outils

Autres langues

Loi log-normale

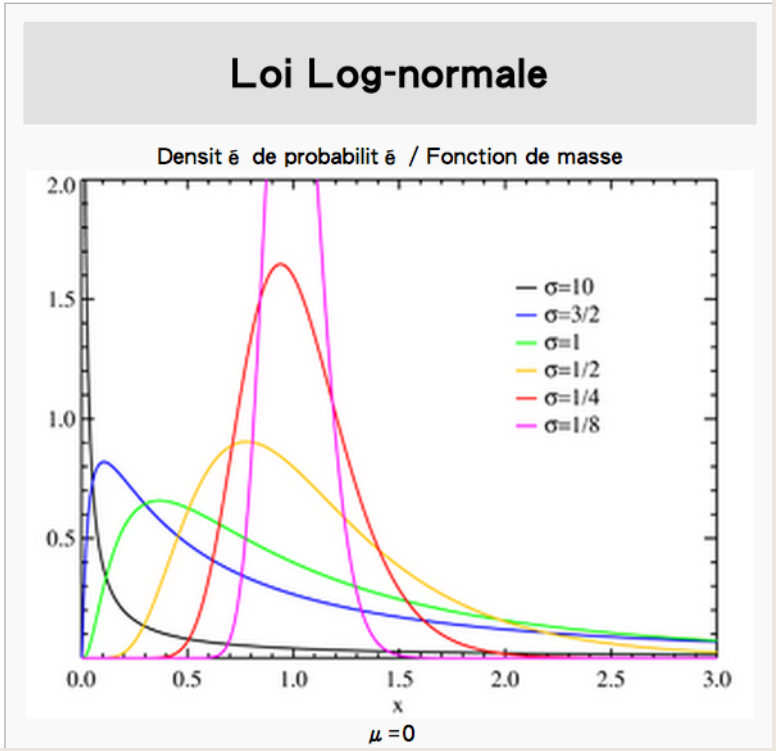
En th é orie des probabilit é s et statistique, une variable al é atoire X est dite suivre une loi log-normale de param è tres μ et σ^2 si la variable $Y = \ln(X)$ suit une loi normale d'esp é rance μ et de variance σ^2.

Cette loi est parfois é galement appel é e loi de Galton.

Une variable peut ê tre mod é lis é e par une loi log-normale si elle est le r é sultat de la multiplication d'un grand nombre de petits facteurs ind é pendants.

Sommaire [masquer]

1 Caract é risation
1.1 Densit é



$x^{c \log x}$ **with** $c < 0$

$\prod_{k\geqslant 0}\left(1+xq^k\right)$ $(0<q<1)$

q-Pochhammer or q-shifted factorial extending to infinity (also #(parts) of partitions into distinct parts)

Asymptotics when $x\to\infty$***?***

$q=\frac{1}{2}$ (often in digital search tree): $\prod_{k\geqslant 0}\left(1+\frac{x}{2^k}\right)\approx$?

$$\prod_{k\geqslant 0}\left(1+\frac{x}{2^k}\right)\approx\prod_{0\leqslant k<\log_2 x}\left(1+\frac{x}{2^k}\right)\approx\prod_{0\leqslant k<\log_2 x}\frac{x}{2^k}$$

$$\approx x^{\log_2 x}2^{-\frac{1}{2}(\log_2 x)^2+\frac{1}{2}\log_2 x}=e^{\frac{(\log x)^2}{2\log 2}}e^{\frac{1}{2}\log x}$$

Mellin transform

$$\sum_{k\geqslant 0}\log\left(1+\frac{z}{2^k}\right)=\frac{(\log z)^2}{2\log 2}+\frac{1}{2}\log z+\underbrace{P(\log_2 z)}_{\text{Periodic function}}+O\left(\frac{1}{|z|}\right)$$

$[z^n] e^{\log^2 \frac{1}{1-z}}$

Saddle-point method (singularity analysis $\Longrightarrow$ polynomial growth)

$$
\begin{aligned}
[z^n] e^{\log^2 \frac{1}{1-z}} &= \frac{1}{2\pi i} \oint_{|z|=r<1} z^{-n-1} e^{\log^2 \frac{1}{1-z}} \mathbf{d}z \\
&= \frac{1}{2\pi i} \oint_{|w|=r<1} (1-w)^{-n-1} e^{(\log w)^2} \mathbf{d}w \\
&\sim \frac{1}{2\pi i} \oint_{|w|=r<1} e^{nw+(\log w)^2} \mathbf{d}w
\end{aligned}
$$

$w = re^{it}$, $\frac{n}{2} = \frac{1}{r} \log \frac{1}{r}$ (Lambert W-function)

$$
\begin{aligned}
[z^n] e^{\log^2 \frac{1}{1-z}} &\sim \frac{e^{nr+(\log r)^2}}{2\pi} \int e^{-\frac{1}{2}(nr+2)t^2} \mathbf{d}t \\
&\sim \frac{n}{2\sqrt{\pi \log n}} \exp\left(\left(\log \frac{n/2}{\log(n/2)} \right)^2 \right)
\end{aligned}
$$

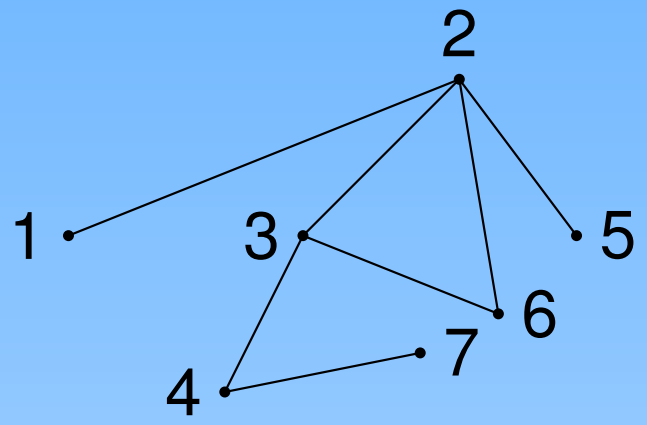
EXHAUSTIVE SEARCH FOR MAXIMUM INDEPENDENT SETS IN RANDOM GRAPHS

& other examples

MAXIMUM INDEPENDENT SET

Independent (stable) set of a graph

A set of vertices no two of which share the same edge.



MIS $= \{1, 3, 5, 7\}$

The MIS (maximum independent set) problem

Search for an independent set with the largest size.

NP hard!!

MAXIMUM INDEPENDENT SET

Equivalent versions

- **MIS $\equiv$ MAXIMUM CLIQUE on the complementary graph (clique $=$ complete subgraph).**
- **MIS $\equiv$ MINIMUM VERTEX COVERING (*every edge connects at least one vertex*) since the complement of a vertex cover in any graph is an independent set.**

Among Karp's (1972) original list of 21 NP-complete problems

Many applications

Computer vision, Pattern recognition, Coding theory, Map labeling, Molecular biology, Chess games, Stock market, . . .

ALTERNATIVE FORMULATIONS

Harant (2000), Abello et al. (2001)

$$|\mathbf{MIS}(G)| = \max_{(x_1,\ldots,x_n)\in[0,1]^n} \left(\sum_{1\leqslant i\leqslant n} x_i - \sum_{(i,j)\in E} x_i x_j \right).$$

***Classical*: Motzkin and Straus (1965, *Canad. J. Math*)**

$$2 \max_{\substack{x_1+\cdots+x_n=1 \\ 0\leqslant x_1,\ldots,x_n\leqslant 1}} \sum_{(i,j)\in E} x_i x_j = 1 - \frac{1}{|\mathbf{max\ clique}|}$$

Widely addressed

***Algorithmic, theoretical and practical* connections of many other formulations. (Ref: Butenko's 2003 PhD Thesis)**

THEORETICAL RESULTS

Random model: Erdős–Rényi's $G_{n,p}$

Vertex set $= \{1, 2, \ldots, n\}$ and all edges occur independently with the same probability p.

The cardinality of an MIS in $G_{n,p}$

Matula (1970), Grimmett and McDiarmid (1975), Bollobás and Erdős (1976), Frieze (1990): If $pn \to \infty$, then

$$|\mathbf{MIS}_n| \sim 2\log_{1/q} pn \quad \textbf{whp},$$

where $q = 1 - p$; and $\exists k = k_n$ such that

$$|\mathbf{MIS}_n| = k \textbf{ or } k+1 \quad \textbf{whp}.$$

$$k_n = 2\log_{1/q} n - 2\log_{1/q}\log_{1/q} n + O(1)$$

A GREEDY ALGORITHM

Adding vertices one after another whenever possible

The size of the resulting IS ($S_0 \equiv 0$)

$$S_n \stackrel{d}{=} 1 + S_{n-1-\mathbf{Binom}(n-1;p)} \qquad (n \geqslant 1).$$

$S_n \equiv |$***right arm of a random digital search tree***$|$

$$\mu_n = 1 + \sum_{0 \leqslant k < n} \binom{n-1}{k} p^k q^{n-1-k} \mu_{n-1-k}$$

$$\tilde{f}'(z) + \tilde{f}(z) = 1 + \tilde{f}(qz), \quad \tilde{f}(z) := e^{-z} \sum_{j \geqslant 0} \frac{\mu_j}{j!} z^n$$

Greedy is $\frac{1}{2}$-optimal (analytic approach): Poisson generating function, Rice's integrals (finite differences)

$$\mathbb{E}(S_n) \sim \log_{1/q} n + \mathbf{Periodic}(\log_{1/q} n).$$

ANALYSIS OF THE GREEDY ALGORITHM

GREEDY IS $\frac{1}{2}$-optimal (probabilistic approach)

Grimmett and McDiarmid (1975), Karp (1976), Fernandez de la Vega (1984), Gazmuri (1984), McDiarmid (1984), ...

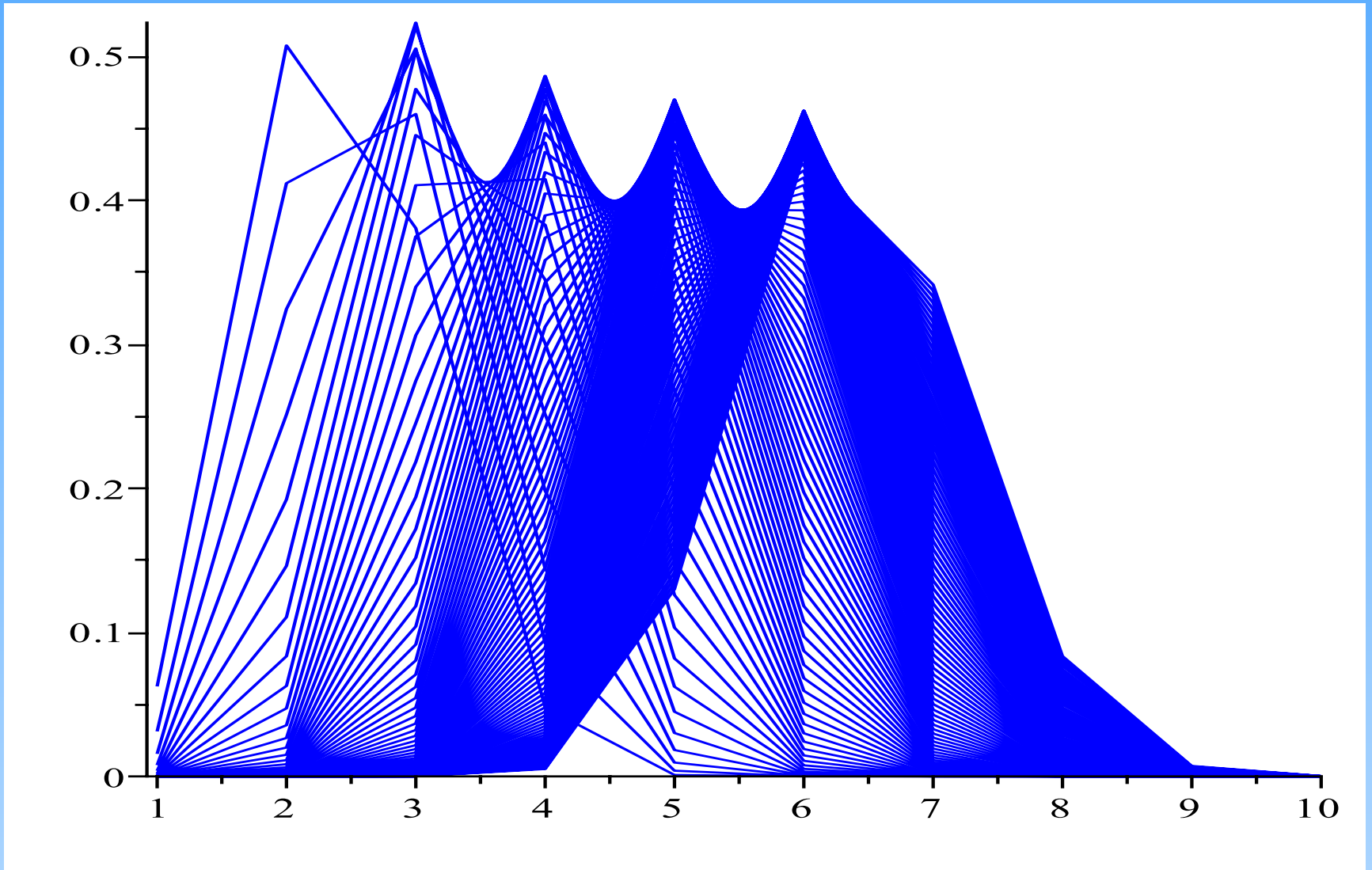
Same analytic approach

- **Variance:** $\mathbb{V}(S_n) \sim$ **a bounded periodic function.**
- **Limit distribution does not exist:** $\mathbb{E}\left(e^{(X_n - \log_{1/q} n)y}\right) \sim F(\log_{1/q} n; y)$**, where**

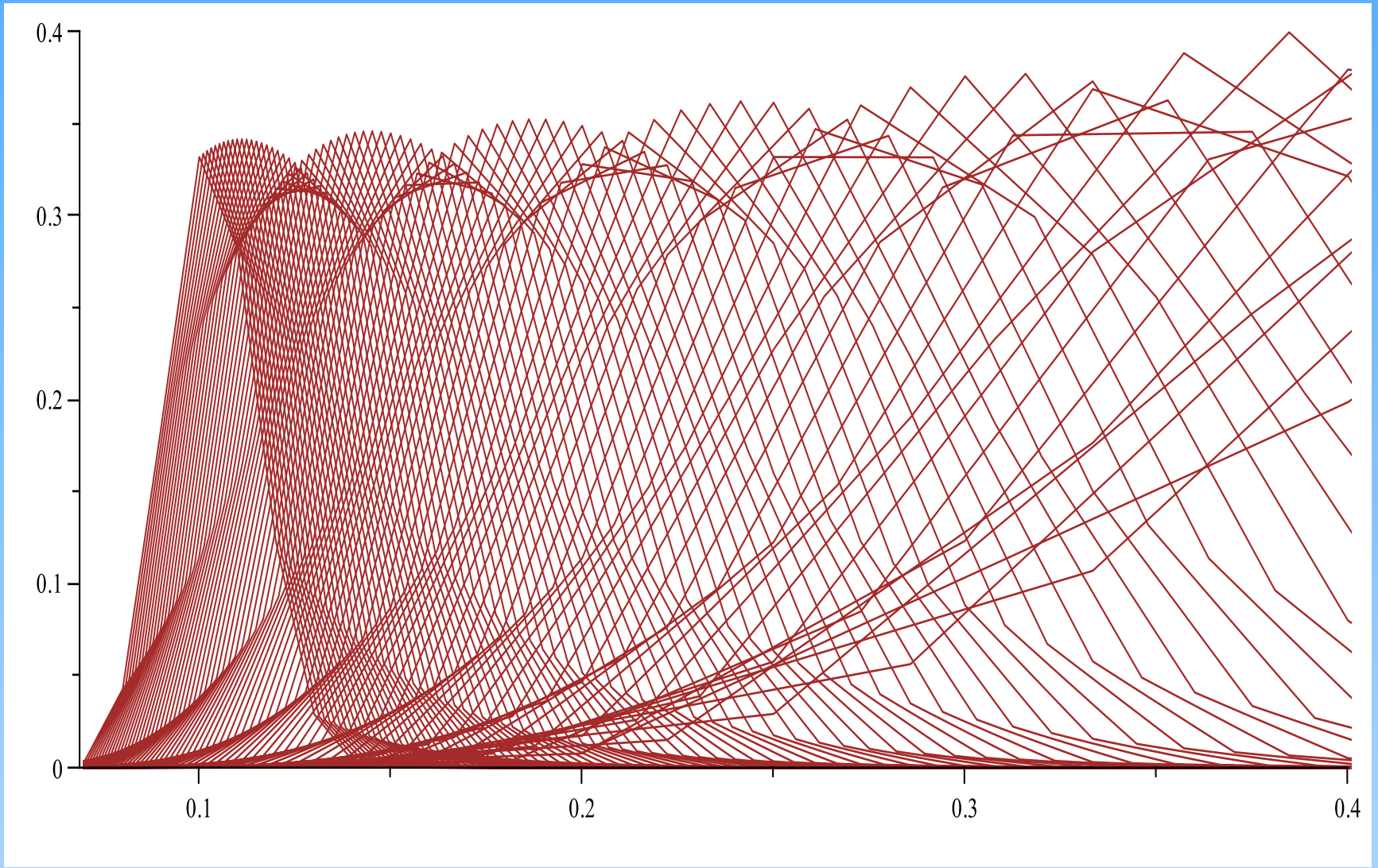
$$F(u;y) := \frac{1-e^y}{\log(1/q)} \left(\prod_{\ell \geqslant 1} \frac{1-e^y q^\ell}{1-q^\ell}\right) \sum_{j \in \mathbb{Z}} \Gamma\left(-\frac{y+2j\pi i}{\log(1/q)}\right) e^{2j\pi i u}.$$

unpublished manuscript

HISTOGRAMS OF S_n FOR $n = 5, \ldots, 100$ ($p = \frac{1}{2}$)



HISTOGRAMS OF S_n FOR $n = 5, \ldots, 100$ ($p = 0.3$)



Normalized to unit interval

A BETTER ALGORITHM?

Jerrum's Metropolis algorithm

He studied a class of algorithms and showed that $\forall \lambda \geqslant 1, \exists$ an initial state from which the expected time for the Metropolis process to reach a clique of size at least $(1+\varepsilon)\log_{1/q}(pn)$ exceeds $n^{\Omega(\log pn)}$.

Frieze and McDiarmid (1997, *RSA*) Research Problem 15

Construct a polynomial time algorithm that finds an independent set of size at least $(\frac{1}{2}+\varepsilon)|MIS_n|$ whp or show that such an algorithm does not exist . . .

Recent advances: Coja-Oghlan and Efthymiou (2011)

Anything beyond $(1+\varepsilon)\log_{1/q} n$ is very difficult.

EXACT AND APPROXIMATE ALGORITHMS

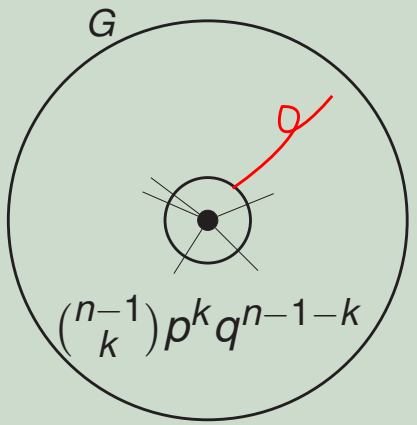
Exact algorithms

A huge number of algorithms proposed in the literature; see Bomze et al.'s survey (in *Handbook of Combinatorial Optimization*, 1999).

A simple exhaustive search algorithm

$$\alpha(G) = \max \left\{ \begin{array}{l} \underbrace{\alpha(G \setminus \{v\})}_{v \notin \mathbf{MIS}(G)}, \\ \underbrace{1 + \alpha(G \setminus N^*(v))}_{v \in \mathbf{MIS}(G)} \end{array} \right\}$$

where $N^*(v)$ denotes the union of v and all its neighbors.



AIM: A MORE PRECISE ANALYSIS OF THE EXHAUSTIVE ALGORITHM

MIS contains either v or not

$$X_n \stackrel{d}{=} X_{n-1} + X^*_{n-1-\mathbf{Binom}(n-1;p)} \qquad (n \geqslant 2),$$

$$(X_0 = 0 \textbf{ and } X_1 = 1)$$

Special cases

- **If $p \sim 1$, then the second term is small, so we expect a *polynomial* time bound.**
- **If $p \sim 0$, then the second term is large, and we expect an *exponential* time bound.**
- **What happens for p in between?**

X_n and X^*_n are strongly dependent

MEAN VALUE

The expected value $\mu_n := \mathbb{E}(X_n)$ satisfies

$$\mu_n = \mu_{n-1} + \sum_{0 \leqslant j < n} \binom{n-1}{j} p^j q^{n-1-j} \mu_{n-1-j} \quad (n \geqslant 2).$$

$$(\mu_0 = 0 \textbf{ and } \mu_1 = 1)$$

Poisson generating function: $\tilde{f}(z) := e^{-z} \sum_{n \geqslant 0} \mu_n z^n / n!$

$$\tilde{f}'(z) = \tilde{f}(qz) + e^{-z}$$

will prove

$$\log \mu_n = \frac{\left(\log \frac{n}{\log_\kappa n}\right)^2}{2 \log \kappa} + \left(\frac{1}{2} + \frac{1}{\log \kappa}\right) \log n - \log \log n$$

$$+ \textbf{Periodic}\left(\log_\kappa \frac{n}{\log n}\right) + o(1) \quad (\kappa := 1/q)$$

EXHAUSTIVE MIS: LITERATURE

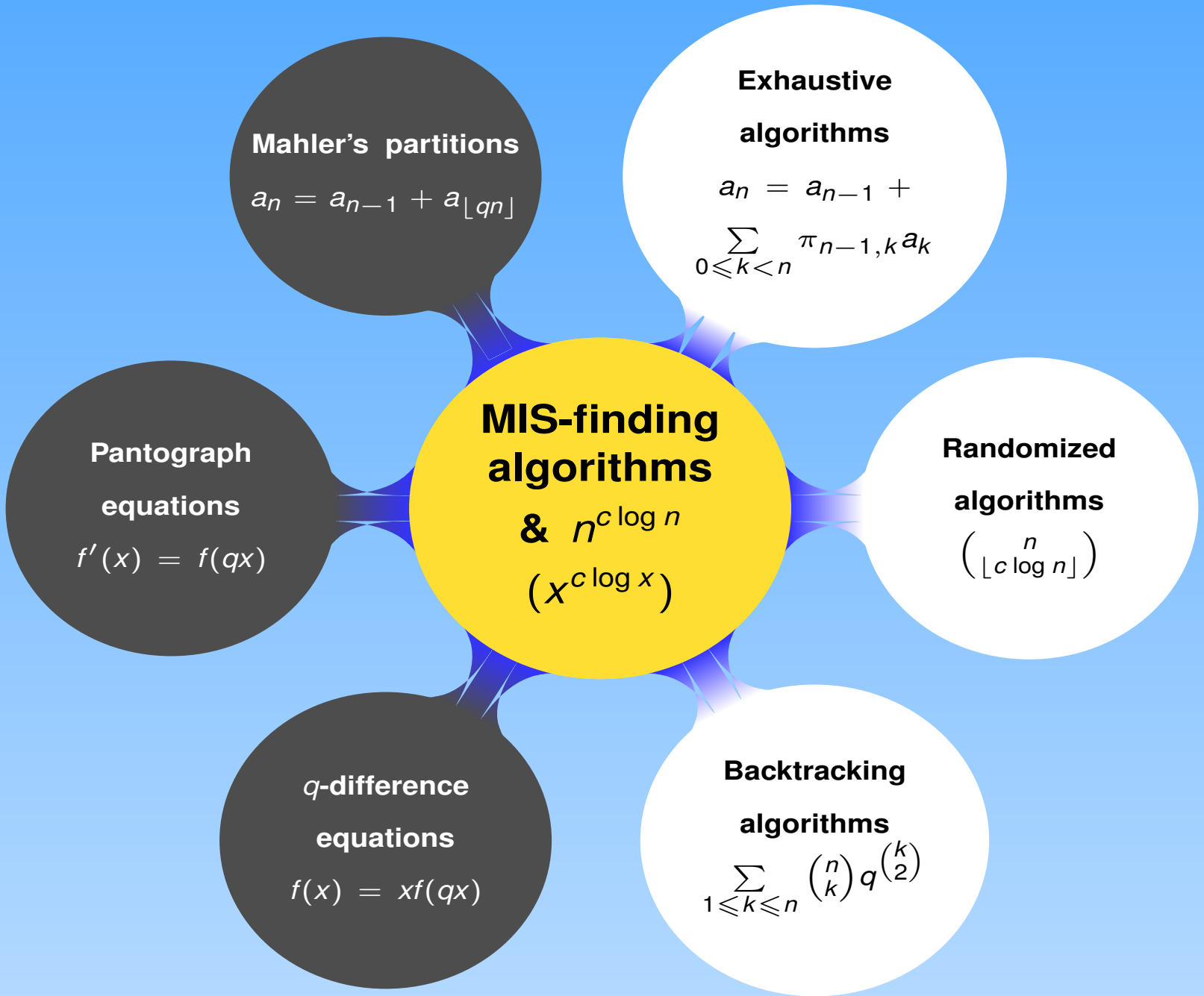
Exhaustive search algorithms

- **Tarjan and Trojanowski (1977): an improved *exhaustive* algorithm of complexity $O(2^{n/3})$.**
- **Chvátal (1977) proposes *exhaustive* algorithms where almost all $G_{n,1/2}$ creates at most $n^{2(1+\log_2 n)}$ subproblems.**
- **Pittel (1982):** ($\kappa := 1/q$)

$$\mathbb{P}\left(n^{\frac{1-\varepsilon}{4}\log_\kappa n} \leqslant \text{Time}^{\text{used by Chvátal's algo}} \leqslant n^{\frac{1+\varepsilon}{2}\log_\kappa n} \right) \geqslant 1 - n^{-c\log n}$$

- **Wilf's (1986) *Algorithms and Complexity* a *backtracking* algorithm enumerating all independent sets with time complexity $n^{O(\log n)}$.**

WHY $n^{c \log n}$?


Mahler's partitions
$a_n = a_{n-1} + a_{\lfloor qn \rfloor}$
Exhaustive algorithms
$a_n = a_{n-1} + \sum_{0 \leqslant k < n} \pi_{n-1,k} a_k$
Pantograph equations
$f'(x) = f(qx)$
MIS-finding algorithms & $n^{c \log n}$ ($x^{c \log x}$)
Randomized algorithms
$\binom{n}{\lfloor c \log n \rfloor}$
q-difference equations
$f(x) = xf(qx)$
Backtracking algorithms
$\sum_{1 \leqslant k \leqslant n} \binom{n}{k} q^{\binom{k}{2}}$

$n^{c \log n}$-ASYMPTOTICS: ALGORITHMS

A simple randomized algorithm

Since $\mathbb{P}(|MIS| = k_n$ **or** $|MIS| = k_n + 1) \to 1$**, check all**

$$\binom{n}{k_n} + \binom{n}{k_n + 1} \leqslant \frac{2}{k!} n^{k_n+1} = O\left(n^{2 \log_\kappa n}\right) \quad \textbf{subsets}.$$

If none is independent ($\mathbb{P}$ very small), use exhaustive.

$n^{c \log n}$-ASYMPTOTICS: MAHLER'S PARTITIONS

Approximate $\mu_n = \mu_{n-1} + \sum_{0 \leqslant j < n} \binom{n-1}{j} p^j q^{n-1-j} \mu_{n-1-j}$ by

$$a_n = a_{n-1} + a_{\lfloor qn \rfloor}.$$

(most mass centers around pn)

Mahler's (1940) partitions (when $\kappa = 1/q \in \mathbb{Z}$)

Number of partitions of κn into positive powers of κ.

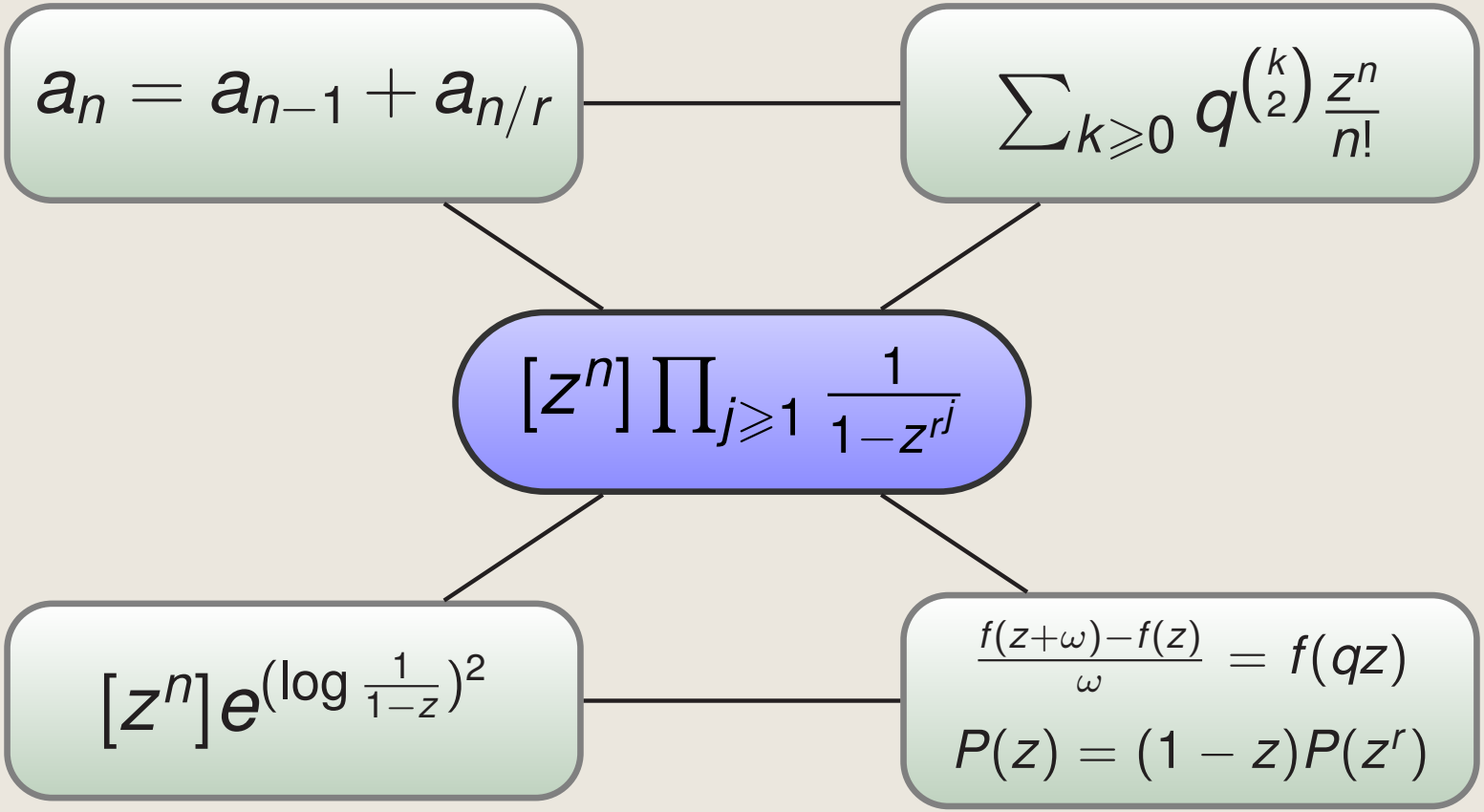
de Bruijn (1948)

$$\log a_n = \frac{\left(\log \frac{n}{\log_\kappa n}\right)^2}{2 \log \kappa} + \left(\frac{1}{2} + \frac{1}{\log \kappa}\right) \log n - \log \log n$$
$$+ \textbf{Periodic}\left(\log_\kappa \frac{n}{\log n}\right) + o(1)$$

$$|\log \mu_n - \log a_n| = \textbf{Periodic}(\cdot) + o(1),$$ **AIMD**

MAHLER'S PARTITIONS

Mahler (1940), de Bruijn (1948)



Jacobi's theta function $\sum_{k\geqslant 0} q^{k^2} z^k$

$n^{c \log n}$-ASYMPTOTICS: PANTOGRAPH EQUATIONS

Continuous version $A(x) - A(x-1) = A(qx)$

Since $A(x) - A(x-1) \approx A'(x)$ for large x, consider

$$A'(x) = A(qx)$$

which is a special case of the so-called *"pantograph equations"*

$$A'(x) = \alpha A(qx) + \beta A(x)$$

$n^{c \log n}$-ASYMPTOTICS: PANTOGRAPH EQUATIONS

Original paper

Ockendon & Tayler (1971, *Proc. Royal Soc. A*) *"The dynamics of a current collection system for an electric locomotive"*

Important papers

- **de Bruijn (1953, *Indagationes Math.*)**
- **Kato & McLeod (1971, *Bull. AMS*)**
- **Iserles (1993, European J. Appl. *Math*)**

$A'(x) = \alpha A(qx)$, $\alpha > 0$ ($\kappa := 1/q$)

de Bruijn (1953) and Kato & McLeod (1971):

$$\log A(x) = \frac{\left(\log \frac{x}{\log_\kappa x}\right)^2}{2 \log \kappa} + \left(\frac{1}{2} + \frac{1}{\log \kappa} + \frac{\log \alpha}{\log \kappa}\right) \log x$$

$$- \left(1 + \frac{\log \alpha}{\log \kappa}\right) \log \log x + \textbf{Periodic}\left(\log_\kappa \frac{x}{\log x}\right) + o(1)$$

$n^{c \log n}$-ASYMPTOTICS: FORMAL ASPECTS

$A'(x) = A(qx)$

Try $A(x) = x^{c \log x}$ (x^α & c^x fail). Then c satisfies

$$x^{1-2c\log\kappa} = 2ce^{c(\log\kappa)^2}\log x$$

Thus $c = \frac{1}{2\log\kappa} + O\left(\frac{\log x}{x}\right)$ and

$$\log A(x) \sim \frac{(\log x)^2}{2\log\kappa}.$$

A formal solution

$$\sum_{j\geqslant 0} \frac{q^{\binom{j}{2}}}{j!} x^j$$

q-difference equation $f(x) = xf(qx)$

A formal solution $\exp\left(\frac{(\log(x/\sqrt{q}))^2}{2\log\kappa}\right)$

$n^{c \log n}$-ASYMPTOTICS: $\mathbb{E}$(# INDEPENDENT SETS)

Expected cost of Wilf's backtracking algorithms:

$$a_n := \sum_{1 \leqslant k \leqslant n} \binom{n}{k} q^{\binom{k}{2}}$$

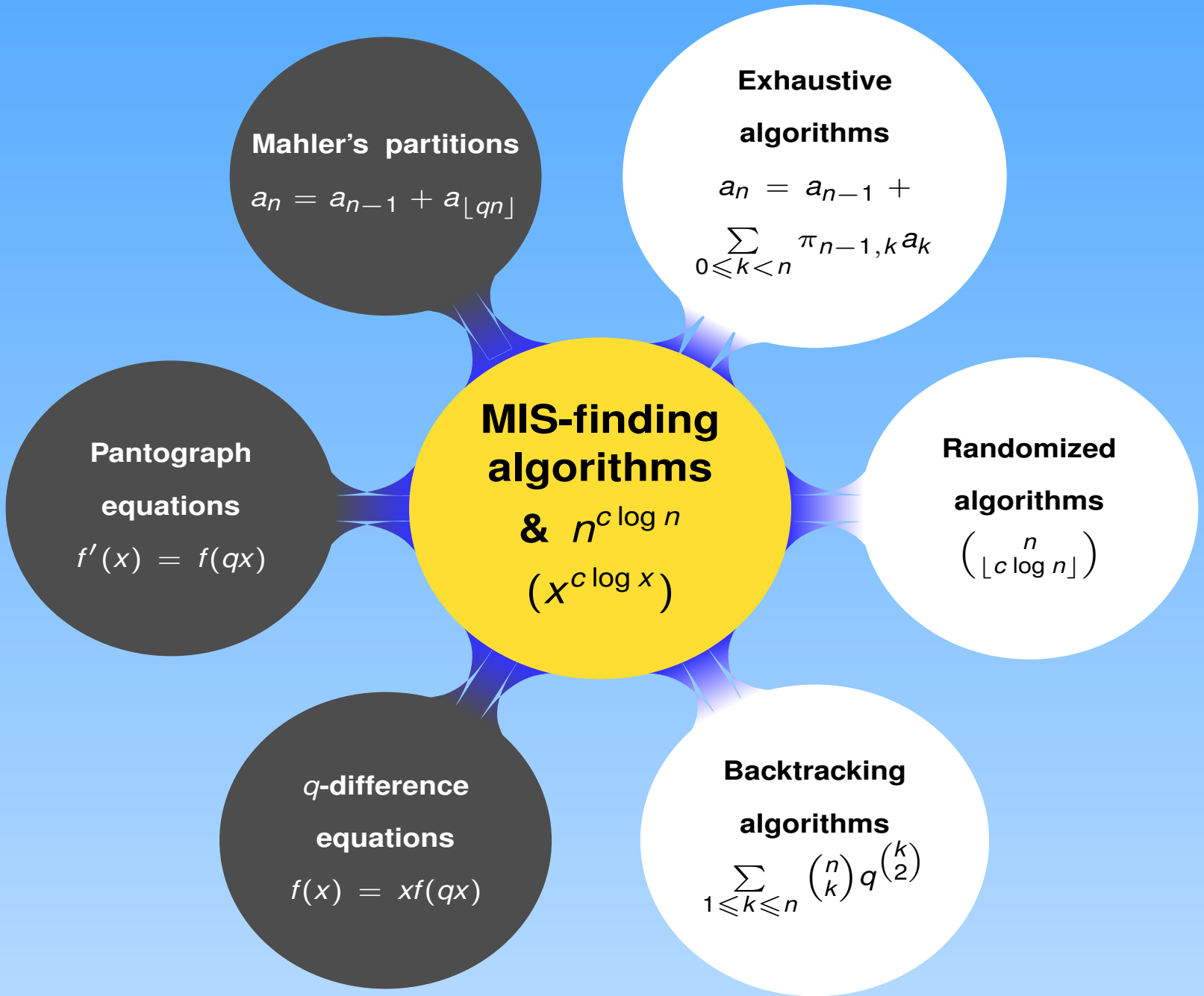
$$a_n = 1 + a_{n-1} + \sum_{0 \leqslant k < n} \binom{n-1}{k} p^k q^{n-1-k} a_{n-1-k}$$

Same pattern of asymptotic behaviors

$$a_n \sim \sum_{k \geqslant 1} \frac{q^{\binom{k}{2}}}{k!} n^k$$

The series $\sum_{k \geqslant 0} \frac{q^{\binom{k}{2}}}{k!} n^k$, first appeared in Mahler (1940), also arises in many applications.

ANALYTIC & ALGORITHMIC SOURCES OF $n^{c \log n}$



$n^{c \log n}$-ASYMPTOTICS: OTHER OCCURRENCES

- **Algorithmics**: isomorphism testing, random digital trees, autocorrelations of strings, information theory, and asymptotics of recurrences;
- **Combinatorics**: partitions into powers, palindromic compositions, combinatorial number theory, and universal tree of minimum complexity;
- **Probability**: log-normal distribution, renewal theory, and total positivity;
- **Algebra**: commutative ring theory, and semigroups;
- **Analysis**: pantograph equations, eigenfunctions of operators, geometric partial differential equations, and q-difference equations.

$n^{c \log n}$: KNUTH (1966)

$$\phi_n = \phi_{n-1} + \phi_{\lfloor n/2 \rfloor} \quad (\phi_0 = 1)$$

In the remainder of this paper we will determine the true rate of growth of the sequence ϕ_n; it will be proved by elementary methods that

$$\ln \phi_n \sim \frac{1}{\ln 4} (\ln n)^2 \quad ,$$

i. e.

$$\phi_n = e^{\frac{1}{\ln 4}(\ln n)^2 + o((\ln n)^2)} \quad . \tag{6}$$

$n^{c \log n}$-ASYMPTOTICS: ALGORITHMS

Miller (1978, *STOC*): Group isomorphisms

On the $n^{\log n}$ Isomorphism Technique*†

A preliminary report

Gary L. Miller
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Tarjan has given an algorithm for deciding isomorphism of two groups of order n (given as multiplication tables) which runs in $O(n^{(\log_2 n + O(1)}$ steps where n is the order of the groups. Tarjan uses the fact that a group of n is generated by log n elements. In this paper, we show that Tarjan's technique

RECENT ADVANCES

Breaking the $n^{\log n}$ Barrier for Solvable-Group Isomorphism

David Rosenbaum
University of Washington,
Department of Computer Science & Engineering
Email: djr@cs.washington.edu

July 7, 2012

Abstract

We consider the *group isomorphism problem*: given two finite groups G and H specified by their multiplication tables, decide if $G \cong H$. The $n^{\log n}$ barrier for group isomorphism has withstood all attacks — even for the special cases of p-groups and solvable groups — ever since the $n^{\log n + O(1)}$ generator-enumeration algorithm. In this work, we present the first significant improvement over $n^{\log n}$ by showing that group isomorphism is $n^{(1/2)\log_p n + O(1)}$ Turing reducible to composition-series isomorphism where p is the smallest prime dividing the order of the group. Combining our reduction with an $n^{O(p/\log p)}$ algorithm for p-group composition-series isomorphism, we obtain an $n^{(1/2)\log n + O(1)}$ algorithm for p-group isomorphism. We then generalize our techniques from p-groups using Sylow bases to derive an $n^{(1/2)\log n + O(\log n/\log\log n)}$ algorithm for solvable-group isomorphism. Finally, we relate group isomorphism to the collision problem which allows us replace the $1/2$ in the exponents with $1/4$ using randomized algorithms and $1/6$ using quantum algorithms.

Population recovery and partial identification

PRP - Results

Facts: μ=0 obliterates all information.

- No polytime algorithm for μ = o(1)

Thm 3 [DRWY] A poly(k, n, ε) algorithm, from **lossy** samples, for every μ > .365…

Thm 4 [WY]: A poly($k^{\log k}$, n, ε) algorithm, from **lossy** and/or **noisy** samples, for every μ > 0

Kearns, Mansour, Ron, Rubinfeld, Schapire, Sellie
exp(k) algorithm for this discrete version
Moitra, Valiant
exp(k) algorithm for Gaussian version
(even when noise is unknown)

$n^{c \log n}$: GUIBAS & Odlyzko (1981, *JCT-A*)

6. Counting the Correlations

In this section we use the recursive predicate Ξ to obtain bounds on the number of distinct correlations of length n.

Theorem 6.1 (the number of correlations). *The number $\kappa(n)$ of distinct correlations of length n satisfies*

$$\left(\frac{1}{2\ln 2}+o(1)\right)\ln^2 n \leqslant \ln \kappa(n) \leqslant \left(\frac{1}{2\ln(3/2)}+o(1)\right)\ln^2 n,$$

as $n \to \infty$.

$n^{c \log n}$: CHUNG & GRAHAM (1981)

OFFPRINTS FROM THE THEORY OF APPLICATIONS OF GRAPHS
EDITED BY Dr. Gary Chartrand

On Trees Containing All Small Trees

F.R.K. CHUNG R.L. GRAHAM*
D. COPPERSMITH

ABSTRACT

In this paper we investigate trees (i.e., connected, acyclic graphs) which contain all trees on n vertices as subgraphs. Let us denote by u(n) the minimum number of vertices such a "universal" tree can have. We prove that:

$$u(n) = n^{(\log n - 2 \log \log n + O(1))/2 \log 2},$$

resolving an earlier conjecture of the authors.

$n^{c \log n}$: Cameron & Erdős (1990)

1.3. All partial sums distinct

Let $f(n)$ be the number of sequences

$$a_1 < \ldots < a_t \leq n$$

for which all sums $\sum \epsilon_i a_i$ $(\epsilon_i = 0,1)$ are distinct.

Proposition 1.6.

$$n^{(1+o(1))\log n/\log 3} \leq f(n) \leq n^{(1+o(1))\log n/\log 2}.$$

$n^{c \log n}$: REZNYKOV & SUSHCHANSKY (2006)

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Journal of Algebra 304 (2006) 712–754

JOURNAL OF Algebra

www.elsevier.com/locate/jalgebra

On the 3-state Mealy automata over an m-symbol alphabet of growth order $[n^{\log n/2\log m}]$

Illya I. Reznykov [a,*], Vitaliy I. Sushchansky [b]

Theorem 2.6. *The growth functions have the following sharp estimates:*

$$\delta_{S_{J_m}}(n) \sim n^{\frac{\log n}{2\log m}};$$

$$\gamma_{J_m}(n) = \gamma_{S_{J_m}}(n) \sim \frac{1}{m}\big(m(n+1)\big)^{\frac{\log(m(n+1))}{2\log m}}.$$

$n^{c \log n}$: DE MARCHIS (2010)

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Journal of Functional Analysis 259 (2010) 2165–2192

JOURNAL OF Functional Analysis

www.elsevier.com/locate/jfa

Generic multiplicity for a scalar field equation on compact surfaces

Francesca De Marchis

$$-\Delta_g u + \frac{\rho}{\int_\Sigma dV_g} = \rho \frac{h(x)e^u}{\int_\Sigma h(x)e^u \, dV_g} \quad x \in \Sigma,\ u \in H_g^1(\Sigma), \qquad (*)$$

For any Σ, as $k \geqslant k_0$, $k_0 \in \mathbb{N}^*$ (independent of Σ),

$$\#\{\textit{solutions of } (*)\} \geqslant C \left(\frac{[\frac{k}{2}]}{\log[\frac{k}{2}]} \right)^{\frac{1}{2l_2} \log(\frac{[\frac{k}{2}]}{\log[\frac{k}{2}]}) + 1 + \frac{ll_2}{l_2}} \left[\frac{k}{2} \right]^{(\frac{1}{l_2} - \frac{1}{2})}, \qquad (1.4)$$

where by $[\frac{k}{2}]$ we mean the integer part of $\frac{k}{2}$, $l_2 := \log 2$ and $ll_2 =: \log\log 2$; so in particular

$n^{c \log n}$: KARLIN & ZIEGLER (1996)

Remarks. *(1) The special choice of* $P(z) = (1+z)^m$ *(where (4.6) is satisfied for all* r*, thus generates a* PF_∞ *sequence), leads to* $p_i = \binom{2m}{i}$*, if* Q_1 *is chosen, and* $p_i = \binom{2m+1}{i}$ *if* Q_2 *is chosen. In this case the inequality (4.2), for* $a > 1$*, is*

$$\sum_{i=0}^{n} \binom{n}{i} a^{(n+2-i)i} < \sum_{i=0}^{n} \binom{n}{i} a^{(n+1-i)(i+1)} . \tag{4.9}$$

For this example, the function

$$f(x) = f_a(x) = \sum_{n=0}^{\infty} \left(a^{-n(n+1)/2} \right) \frac{x^n}{n!} \tag{4.10}$$

WHY PRECISE ASYMPTOTIC EXPANSION?

$$\log \mu_n = \frac{\left(\log \frac{n}{\log_\kappa n}\right)^2}{2 \log \kappa} + \left(\frac{1}{2} + \frac{1}{\log \kappa}\right) \log n - \log \log n$$
$$+ \textbf{Periodic}\left(\log_\kappa \frac{n}{\log n}\right) + o(1) \quad (\kappa := 1/q)$$

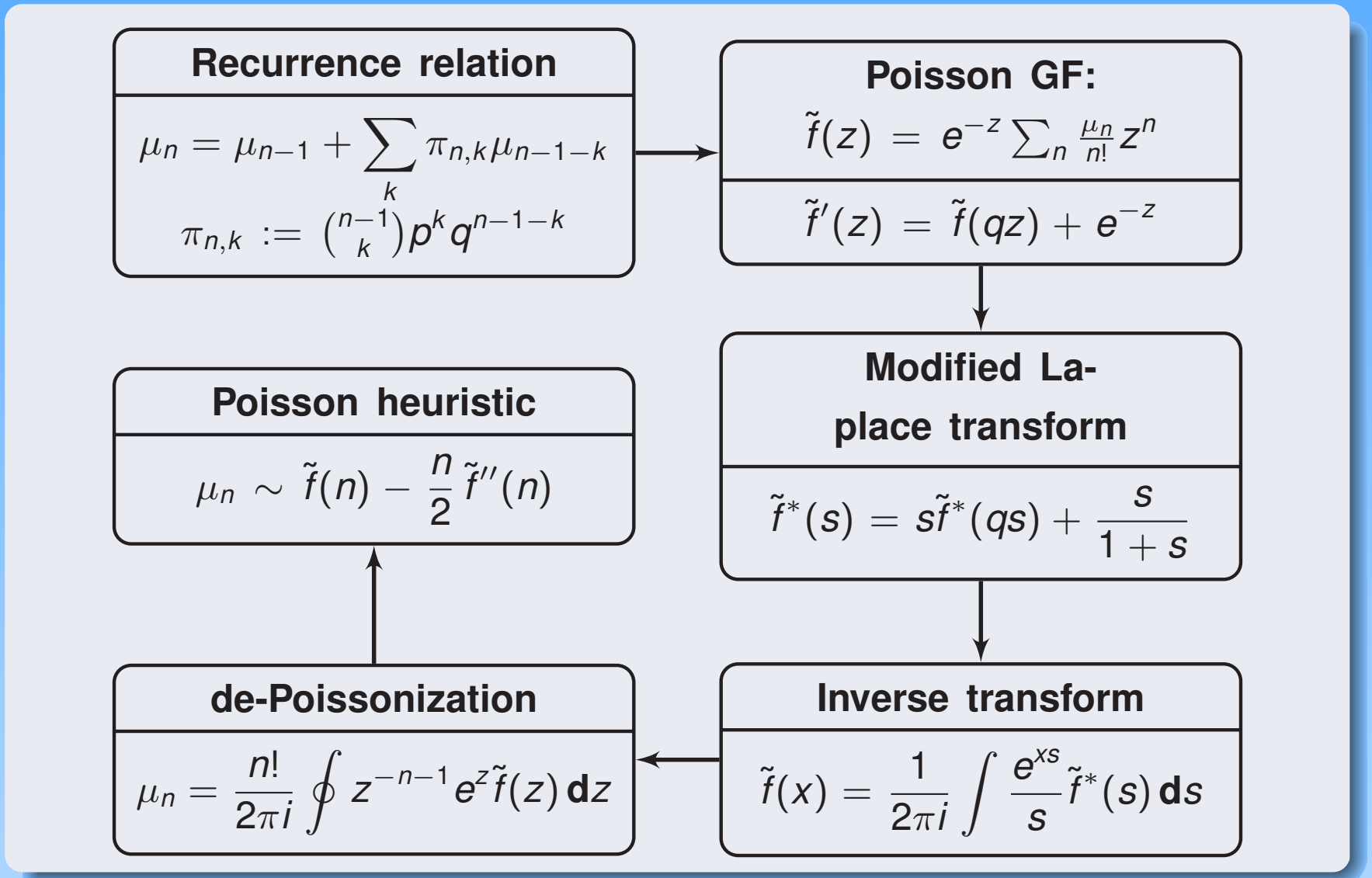
Some reasons

- **Asymptotics: $n^{c \log n}$ is less precise**

$$\frac{\mu_n}{n^{\frac{1}{2}\log_\kappa n}} \asymp \frac{(\log n)^{\frac{1}{2}\log_\kappa \log n - 1 - \frac{\log\log\kappa}{\log\kappa}}}{n^{\log_\kappa \log n - \frac{1}{2} - \frac{1}{\log\kappa} - \frac{\log\log\kappa}{\log\kappa}}} \ll n^{-K} \to 0,$$

- **Numerical "opacity"**
- **Methodology: new tools developed**
- **Generality: discovered**

OUR ANALYTIC APPROACH



RESOLUTION OF THE RECURRENCE

$$\tilde{f}'(z) = \tilde{f}(qz) + e^{-z}$$

The Laplace transform of $\tilde{f}$

$$\mathscr{L}(s) = \int_0^\infty e^{-xs}\tilde{f}(x)\,\mathrm{d}x$$

satisfies

$$s\mathscr{L}(s) = \frac{1}{q}\mathscr{L}\left(\frac{s}{q}\right) + \frac{1}{s+1}.$$

Exact solutions

$$\mathscr{L}(s) = \sum_{j\geqslant 0} \frac{q^{\binom{j+1}{2}}}{s^{j+1}(s+q^j)}.$$

Hard part: asymptotics of $\mathscr{L}(s)$ as $s \to 0$ in $\mathbf{C}$

RESOLUTION OF THE RECURRENCE

Exact solutions

$$\mathscr{L}(s) = \sum_{j\geqslant 0} \frac{q^{\binom{j+1}{2}}}{s^{j+1}(s+q^j)}.$$

Inverting gives $\tilde{f}(z) = \sum_{j\geqslant 0} \frac{q^{\binom{j+1}{2}}}{j!} z^{j+1} \int_0^1 e^{-q^j uz}(1-u)^j \,\mathrm{d}u$**.**

Thus $\mu_n = \sum_{1\leqslant j\leqslant n} \binom{n}{j} (-1)^j \sum_{1\leqslant \ell\leqslant j} (-1)^\ell q^{j(\ell-1)-\binom{\ell}{2}}$**, or**

$$\mu_n = n \sum_{0\leqslant j<n} \binom{n-1}{j} q^{\binom{j+1}{2}} \sum_{0\leqslant \ell<n-j} \binom{n-1-j}{\ell} \frac{q^{j\ell}(1-q^j)^{n-1-j-\ell}}{j+\ell+1}.$$

Neither is useful for numerical purposes for large n**.**

A MORE PRECISE EXPANSION FOR $\tilde{f}(x)$

Method of proof: functional eq. + saddle-point

$$\tilde{f}(x) = \frac{1}{2\pi i} \int_{1-i\infty}^{1+i\infty} e^{xs} \mathscr{L}(s)\, \mathrm{d}s$$

Asymptotics of $\tilde{f}(x)$

Let $\mathit{kappa} = 1/q$ **and** $R \log R = x \log \kappa$**. Then**

$$\tilde{f}(x) \sim \frac{R^{1/\log\kappa + 1/2} e^{(\log R)^2/(2\log\kappa)} G(\log_\kappa R)}{\sqrt{2\pi \log_\kappa R}} \left(1 + \sum_{j\geqslant 1} \frac{\phi_j(\log_\kappa R)}{(\log_\kappa R)^j}\right),$$

as $x \to \infty$**, where** $G(u) := q^{(\{u\}^2+\{u\})/2} F(q^{-\{u\}})$**,**

$$F(s) = \sum_{-\infty<j<\infty} \frac{q^{j(j+1)/2}}{1+q^j s}\, s^{j+1},$$

and the $\phi_j(u)$**'s are bounded,** 1**-periodic functions of** u**.**

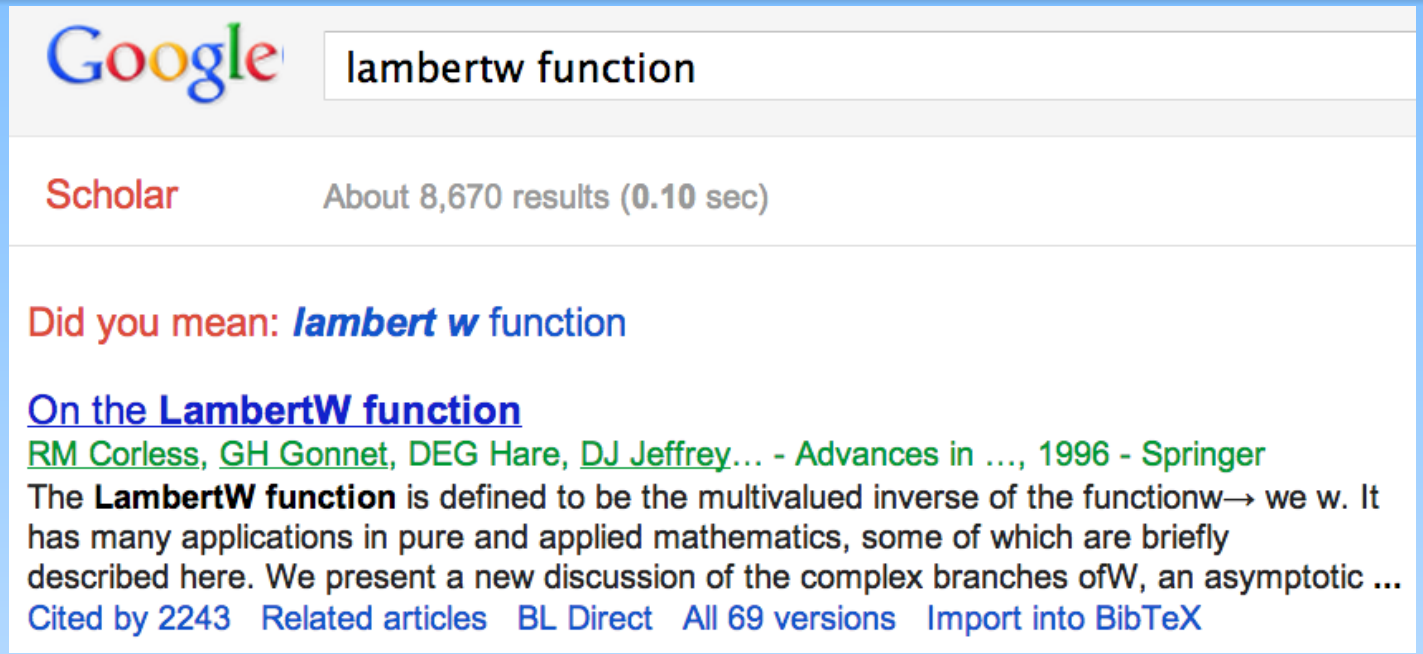
A MORE EXPLICIT ASYMPTOTIC APPROXIMATION

$R = x \log \kappa W(x \log \kappa)$, Lambert's W-function: $W(x)e^{W(x)} = x$

$$W(x) = \log x - \log\log x + \frac{\log\log x}{\log x} + \frac{(\log\log x)^2 - 2\log\log x}{2(\log x)^2} + \cdots .$$

So that

$$\tilde{f}(x) \sim \frac{G\left(\log_\kappa \frac{x}{\log_\kappa x}\right)}{\sqrt{2\pi}} \cdot \frac{x^{1/\log\kappa + 1/2}}{\log_\kappa x} \exp\left(\frac{\left(\log \frac{x}{\log_\kappa x}\right)^2}{2\log\kappa}\right).$$



JUSTIFICATION OF THE POISSON HEURISTIC

$$\mu_n = \frac{n!}{2\pi i} \oint_{|z|=n} z^{-n-1} e^z \tilde{f}(z)\, dz$$

$$\mu_n \sim \frac{G\left(\log_\kappa \frac{n}{\log_\kappa n}\right)}{\sqrt{2\pi}} \cdot \frac{n^{1/\log\kappa + 1/2}}{\log_\kappa n} \exp\left(\frac{\left(\log \frac{n}{\log_\kappa n}\right)^2}{2\log\kappa}\right)$$

Four properties are sufficient: $\mu_n \sim \tilde{f}(n)$

- $\tilde{f}'(z) = \tilde{f}(qz) + e^{-z}$;
- $F(s) = sF(qs)$ $\left(F(s) = \sum_{j\in\mathbb{Z}} \frac{q^{j(j+1)/2} s^{j+1}}{1+q^j s}\right)$;
- $\frac{\tilde{f}^{(j)}(x)}{\tilde{f}(x)} \sim \left(\frac{\log_\kappa x}{x}\right)^j$;
- $|f(z)| \leqslant f(|z|)$ **where** $f(z) := e^z \tilde{f}(z)$.

EXHAUSTIVE ALGORITHM: BEYOND THE MEAN

Intractable due to strong dependence

$$X_n \stackrel{d}{=} \underbrace{X_{n-1} + X^*_{n-1-\mathbf{Binom}(n-1;p)}}_{\textbf{dependent}}$$

Idealized: $Y_n \stackrel{d}{=} \underbrace{Y_{n-1} + Y^*_{n-1-\mathrm{Binom}(n-1;p)}}_{\textit{independent}}$

$$M_n(t) = M_{n-1}(t) \sum_{0 \leqslant j < n} \underbrace{\binom{n-1}{j} p^j q^{n-1-j}}_{\pi_{n,j}} M_{n-1-j}(t)$$

$\bar{M}_n(t) := \mathbb{E}(e^{(Y_n - \mu_n)t})$ $(\Delta_{n,j} := \mu_j + \mu_{n-1} - \mu_n)$

$$\bar{M}_n(t) = \bar{M}_{n-1}(t) \sum_{0 \leqslant j < n} \pi_{n,j} \bar{M}_{n-1-j}(y) e^{\Delta_{n,j} t}$$

ASYMPTOTIC NORMALITY OF Y_n

Convergence in distribution

The distribution of Y_n is asymptotically normal

$$\frac{Y_n - \mu_n}{\sigma_n} \xrightarrow{d} \mathcal{N}(0,1),$$

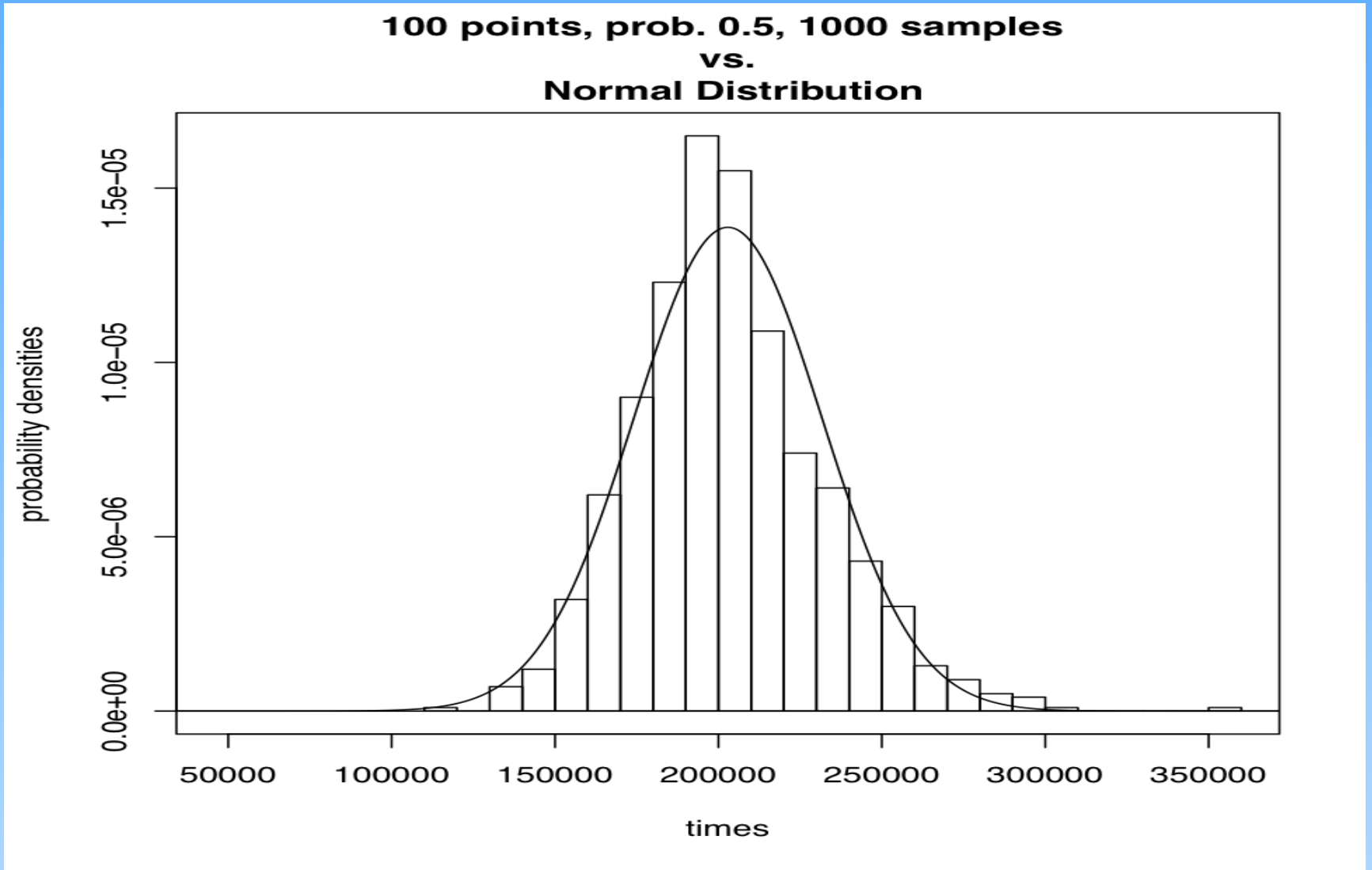
with convergence of all moments.

Proof by the method of moments

- **Derive recurrence for $\mathbb{E}(Y_n - \mu_n)^m$.**
- **Prove by induction (using the asymptotic transfer) that**

$$\mathbb{E}(Y_n - \mu_n)^m \begin{cases} \sim \dfrac{(m)!}{(m/2)!2^{m/2}}\,\sigma_n^m, & \text{if } 2 \mid m, \\ = o(\sigma_n^m), & \text{if } 2 \nmid m, \end{cases}$$

HISTOGRAMS OF X_n (DEPENDENT) FOR $n = 100$ ($p = \frac{1}{2}$)



A NATURAL VARIANT: BINOMIAL $\longrightarrow$ UNIFORM

What happens if $Z_n \stackrel{d}{=} Z_{n-1} + Z^*_{\text{uniform}[0,n-1]}$?

$$\mu_n = \mu_{n-1} + \frac{1}{n} \sum_{0 \leqslant j < n} \mu_j,$$

satisfies $\mu_n \sim c n^{-1/4} e^{2\sqrt{n}}$. **Note:** $\mu_n \approx \mu_{n-1} + \mu_{n/2}$ **fails.**

$$c := \frac{1}{2}\sqrt{\frac{e}{\pi}} \left(e^{-1} - \int_1^\infty v^{-1} e^{-v} \, \mathrm{d}v \right) \approx 0.06906\,46192\ldots$$

Limit law not Gaussian (by method of moments)

$$\frac{Z_n}{\mu_n} \xrightarrow{d} Z,$$

where $g(z) := \sum_{m \geqslant 1} \mathbb{E}(Z^m) z^m / (m \cdot m!)$ **satisfies**

$$z^2 g'' + z g' - g = z g g'.$$

AN INTERESTING PROPERTY

By the psi-series method for nonlinear DEs

$$\mathbb{E}(Z^m) = m \cdot m! \rho^{-m} \left(2 + \frac{2}{3\,m^2} + O\left(m^{-3}\right) \right),$$

(Chern, Fernández-Camacho, Hwang, Martínez, *RSA*)

Original motivation: $p_n = n^{-2} \sum_{0 \leqslant j < n} p_j p_{n-1-j}$ $(p_0 = 1)$

Equality of two random binary search trees

$$zP'' + P' = P^2$$

$$p_n = \rho^{-n-1} \left(6n + \frac{18}{5} + \frac{336}{3125\,n^5} + O\left(n^{-6}\right) \right)$$

Applied to > 10 ***structures***

CONCLUSION

$n^{\log n}$ est ainsi une échelle

- **au carrefour de nombreux problèmes**
- **signature intrinsèque de récurrences ”mixtes” (additif/multiplicatif) et des mathématiques associées**
- **connecte des structures à priori diverses**
- **oblige à développer de nouvelles mathématiques**
- **riche domaine pour la combinatoire (graphes, analytique)**
- **améliore notre compréhension des paysages de complexité de problèmes (NP-complets ou pas)**