Phase transitions of Random Constraint Satisfaction Problems

Allan Sly, Princeton University

ICM, August 2018
Introduction:
random constraint satisfaction problems;
Constraint satisfaction problem (CSP): is it possible to assign values to a set of variables to satisfy a given set of constraints?

Scheduling your appointments for the day

System of linear equations.

Colouring a graph or finding a large independent set.

Satisfying a Boolean formula.

Random CSPs

Our focus is to investigate properties when the constraints are chosen randomly.
Combinatorics and Theoretical Computer Science

Constraint satisfaction problem (CSP): is it possible to assign values to a set of variables to satisfy a given set of constraints?

- Scheduling your appointments for the day
- System of linear equations.
- Colouring a graph or finding a large independent set.
- Satisfying a Boolean formula.
Combinatorics and Theoretical Computer Science

Constraint satisfaction problem (CSP): is it possible to assign values to a set of *variables* to satisfy a given set of *constraints*?

- Scheduling your appointments for the day
- System of linear equations.
- Colouring a graph or finding a large independent set.
- Satisfying a Boolean formula.

Random CSPs

Our focus is to investigate properties when the constraints are chosen randomly.
Combinatorial properties of Random Graphs:

- Erdős-Rényi Random Graph: $G(n, \alpha/n)$ with $n$ vertices and edges with probability $\alpha/n$ (average degree $\alpha$).
- Random $\alpha$-regular graph: Uniformly chosen from $\alpha$-regular graphs on $n$ vertices.
Combinatorial properties of Random Graphs:

- Erdős-Rényi Random Graph: $G(n, \alpha/n)$ with $n$ vertices and edges with probability $\alpha/n$ (average degree $\alpha$).
- Random $\alpha$-regular graph: Uniformly chosen from $\alpha$-regular graphs on $n$ vertices.

When is there a random proper $k$-colouring?
Combinatorial properties of Random Graphs:

- Erdős-Rényi Random Graph: $G(n, \alpha/n)$ with $n$ vertices and edges with probability $\alpha/n$ (average degree $\alpha$).
- Random $\alpha$-regular graph: Uniformly chosen from $\alpha$-regular graphs on $n$ vertices.

When is there a random proper $k$-colouring?

Is there an independent set of size $\beta n$?
The random K-SAT problem, a model of a random Boolean formula, is perhaps the canonical random CSP.
**K-SAT** The random K-SAT problem, a model of a random Boolean formula, is perhaps the canonical random CSP.

**Basic Definition:**

Variables: $x_1, \ldots, x_n \in \{\text{TRUE, FALSE}\} \equiv \{+, -\}$

Constraints: $m$ clauses taking the OR of $k$ variables uniformly chosen from $\{+x_1, -x_1, \ldots, +x_n, -x_n\}$.
**K-SAT** The random K-SAT problem, a model of a random Boolean formula, is perhaps the canonical random CSP.

**Basic Definition:**

Variables: $x_1, \ldots, x_n \in \{\text{TRUE, FALSE}\} \equiv \{+, -\}$

Constraints: $m$ clauses taking the OR of $k$ variables uniformly chosen from $\{+x_1, -x_1, \ldots, +x_n, -x_n\}$.

Example: A 3-SAT formula with 4 clauses:

$$G(x) = ( +x_1 \text{ OR } +x_2 \text{ OR } -x_3 ) \text{ AND } ( +x_3 \text{ OR } +x_4 \text{ OR } -x_5 )$$

$$\text{AND} ( -x_1 \text{ OR } -x_4 \text{ OR } +x_5 ) \text{ AND } ( +x_2 \text{ OR } -x_3 \text{ OR } +x_4 )$$
**K-SAT** The random K-SAT problem, a model of a random Boolean formula, is perhaps the canonical random CSP.

**Basic Definition:**

Variables: \( x_1, \ldots, x_n \in \{ \text{TRUE}, \text{FALSE} \} \equiv \{+, -\} \)

Constraints: \( m \) clauses taking the OR of \( k \) variables uniformly chosen from \( \{+x_1, -x_1, \ldots, +x_n, -x_n\} \).

Example: A 3-SAT formula with 4 clauses:

\[
\mathcal{G}(x) = (+x_1 \text{ OR } +x_2 \text{ OR } -x_3) \text{ AND } (+x_3 \text{ OR } +x_4 \text{ OR } -x_5) \\
\text{ AND } (-x_1 \text{ OR } -x_4 \text{ OR } +x_5) \text{ AND } (+x_2 \text{ OR } -x_3 \text{ OR } +x_4)
\]

**Clause density:** The K-SAT model is parameterized the problem by the density of clauses \( \alpha = m/n \).
**K-SAT** The random K-SAT problem, a model of a random Boolean formula, is perhaps the canonical random CSP.

**Basic Definition:**

Variables: $x_1, \ldots, x_n \in \{\text{TRUE}, \text{FALSE}\} \equiv \{+, -\}$

Constraints: $m$ clauses taking the OR of $k$ variables uniformly chosen from $\{+x_1, -x_1, \ldots, +x_n, -x_n\}$.

Example: A 3-SAT formula with 4 clauses:

$$G(x) = (+x_1 \text{ OR } +x_2 \text{ OR } -x_3) \text{ AND } (+x_3 \text{ OR } +x_4 \text{ OR } -x_5)$$

$$\text{AND } (-x_1 \text{ OR } -x_4 \text{ OR } +x_5) \text{ AND } (+x_2 \text{ OR } -x_3 \text{ OR } +x_4)$$

**Clause density:** The K-SAT model is parameterized the problem by the density of clauses $\alpha = m/n$.

**Variant NAE-SAT:** An assignment $\overline{x}$ is a solution if both $\overline{x}$ and $\overline{-x}$ are satisfying.
**Graphical description**: We can encode a K-SAT formula as a bipartite hyper-graph:

Take a 4-SAT formula with 3 clauses:

\[ p \lor x_1 \lor x_3 \lor -x_5 \lor -x_7 q \land p \lor -x_1 \lor -x_2 \lor +x_5 \lor +x_6 q \]

We can encode the formula as a bipartite graph \( G = (V, F, E) \):

- **variables** \( V \)
- **clauses** \( F \)
- **edges** \( E \)

Clause \( a \) P \( F \), variable \( v \) P \( V \):

- **blue edge** \((av)\) if \(+x_v\) in clause \( a \)
- **yellow edge** \((av)\) if \(-x_v\) in clause \( a \)

(4-SAT: each clause has degree 4)

The resulting random graph is locally tree-like, almost no short cycles and it's local distribution can be described completely.
**Graphical description:** We can encode a K-SAT formula as a bipartite hyper-graph:

Take a 4-SAT formula with 3 clauses: \( G(x) = \)

\[
( +x_1 \text{ OR } +x_3 \text{ OR } -x_5 \text{ OR } -x_7 ) \text{ AND } ( -x_1 \text{ OR } -x_2 \text{ OR } +x_5 \text{ OR } +x_6 ) \]

\[
\text{AND } ( -x_3 \text{ OR } +x_4 \text{ OR } -x_6 \text{ OR } +x_7 )
\]
**Graphical description:** We can encode a K-SAT formula as a bipartite hyper-graph:

Take a 4-SAT formula with 3 clauses: \( \mathcal{G}(\overline{x}) = (\overline{+x_1 OR +x_3 OR -x_5 OR -x_7}) \AND (\overline{-x_1 OR -x_2 OR +x_5 OR +x_6}) \AND (\overline{-x_3 OR +x_4 OR -x_6 OR +x_7}) \)

We can encode the formula as a bipartite graph \( \mathcal{G} \equiv (V, F, E) \):

(4-SAT: each clause has degree 4)
**Graphical description**: We can encode a $K$-SAT formula as a bipartite hyper-graph:

Take a 4-SAT formula with 3 clauses:

$G(x) = ( +x_1 \ OR \ +x_3 \ OR \ -x_5 \ OR \ -x_7 ) \ AND ( -x_1 \ OR \ -x_2 \ OR \ +x_5 \ OR \ +x_6 ) \ AND ( -x_3 \ OR \ +x_4 \ OR \ -x_6 \ OR \ +x_7 )$

We can encode the formula as a bipartite graph $\mathcal{G} \equiv (V, F, E)$:

The resulting random graph is locally tree-like, almost no short cycles and it’s local distribution can be described completely.
Main Question: Satisfiability Threshold
For which $\alpha$ are there satisfying assignments?

Other Questions:
Free Energy: How many solutions are there?
Local Statistics: Properties of solutions such as how many clauses are satisfied only once?
Algorithmic: Can solutions be found efficiently?
Main Question:

- Satisfiability Threshold: For which $\alpha$ are there satisfying assignments?
Main Question:

- *Satisfiability Threshold*: For which $\alpha$ are there satisfying assignments?

Other Question:

- *Free Energy*: How many solutions are there?
- *Local Statistics*: Properties of solutions such as how many clauses are satisfied only once?
- *Algorithmic*: Can solutions be found efficiently?
The Satisfiability Conjecture.

For each $k \geq 2$, random $k$-SAT has a sharp satisfiability threshold $\alpha_{sat}$.

Increasing $\alpha$ $p_{SAT} \rightarrow 1$ as $n \rightarrow \infty$ $p_{UNSAT} \rightarrow 1$ as $n \rightarrow \infty$

with $k$ fixed — that is, a single critical value $\alpha_{sat}$ separates SAT $|$ UNSAT (with high probability in the limit $n \rightarrow \infty$; fixed $k$).

For general $k$, Friedgut ('99) proved the transition sharpens around a (possibly non-convergent) threshold sequence $\alpha_{sat}(n)$ (whereas conjecture requires $\alpha_{sat} \rightarrow \alpha_{sat}$ as $n \rightarrow \infty$).
The Satisfiability Conjecture. For each $k \geq 2$, random $k$-SAT has a sharp satisfiability threshold $\alpha_{\text{sat}}$. 
The Satisfiability Conjecture. For each $k \geq 2$, random $k$-SAT has a sharp satisfiability threshold $\alpha_{\text{sat}}$.

With $k$ fixed, $P(\text{SAT})$ with $n = 12$ decreases as $\alpha$ increases.
The Satisfiability Conjecture. For each \( k \geq 2 \), random \( k \)-SAT has a sharp satisfiability threshold \( \alpha_{\text{sat}} \).

- \( \mathbb{P}(\text{SAT}) \) with \( n = 13 \)

with \( k \) fixed
The Satisfiability Conjecture. For each $k \geq 2$, random $k$-SAT has a sharp satisfiability threshold $\alpha_{\text{sat}}$. 

$\mathbb{P}(\text{SAT})$ with $n = 14$

Increasing $\alpha$ with $k$ fixed
**The Satisfiability Conjecture.** For each $k \geq 2$, random $k$-SAT has a sharp satisfiability threshold $\alpha_{\text{sat}}$. \\

*with $n = 15$* \\

$P(\text{SAT})$ \\

\[\frac{\text{increasing } \alpha}{\text{with } k \text{ fixed}}\]
The Satisfiability Conjecture. For each $k \geq 2$, random $k$-SAT has a sharp satisfiability threshold $\alpha_{\text{sat}}$.

with $k$ fixed

$\mathbb{P}($SAT$)$

with $n = 16$

increasing $\alpha$
The Satisfiability Conjecture. For each $k \geq 2$, random $k$-SAT has a sharp satisfiability threshold $\alpha_{\text{sat}}$. 

$\mathbb{P}(\text{SAT})$ with $n = 17$
The Satisfiability Conjecture. For each $k \geq 2$, random $k$-SAT has a sharp satisfiability threshold $\alpha_{\text{sat}}$. 

\[ \mathbb{P}(\text{SAT}) \]
with $n = 18$

with $k$ fixed

increasing $\alpha$
The Satisfiability Conjecture. For each $k \geq 2$, random $k$-SAT has a sharp satisfiability threshold $\alpha_{\text{sat}}$. 

With $k$ fixed, $\mathbb{P}(\text{SAT})$ with $n = 19$. 

Increasing $\alpha$. 

With $k$ fixed.
The Satisfiability Conjecture. For each $k \geq 2$, random $k$-SAT has a sharp satisfiability threshold $\alpha_{\text{sat}}$. 

$\mathbb{P}(\text{SAT})$ with $n = 20$

with $k$ fixed
The Satisfiability Conjecture. For each \( k \geq 2 \), random \( k\text{-SAT} \) has a sharp satisfiability threshold \( \alpha_{\text{sat}} \).

*with \( k \) fixed*
The Satisfiability Conjecture. For each $k \geq 2$, random $k$-SAT has a sharp satisfiability threshold $\alpha_{\text{sat}}$. 

That is, a single critical value $\alpha_{\text{sat}}$ separates SAT|UNSAT (with high probability in the limit $n \to \infty$; fixed $k$)
The Satisfiability Conjecture. For each $k \geq 2$, random $k$-SAT has a sharp satisfiability threshold $\alpha_{\text{sat}}$.

— that is, a single critical value $\alpha_{\text{sat}}$ separates SAT|UNSAT (with high probability in the limit $n \to \infty$; fixed $k$)

For general $k$, Friedgut (’99) proved the transition sharpens around a (possibly non-convergent) \textit{threshold sequence} $\alpha_{\text{sat}}(n)$ (whereas conjecture requires $\alpha_{\text{sat}}(n) \to \alpha_{\text{sat}}$ as $n \to \infty$)
**Theoretical Physics**

Disordered systems such as *spin glasses* are models of interacting particles/variables with frustrated interactions. Many random constraint satisfaction problems can be recast as dilute mean-field spin glasses.

Increasing constraint density

![Diagram showing increasing constraint density](image)
Theoretical Physics
Disordered systems such as spin glasses are models of interacting particles/variables with frustrated interactions. Many random constraint satisfaction problems can be recast as dilute mean-field spin glasses.

One-step Replica Symmetry Breaking Predictions:
Developed to study dense spin-glasses such as the Sherrington-Kirkpatrick model.

- **Replica Symmetry Breaking**: Clustering of assignments.
- **Cavity Method**: Heuristic for analyzing adding one variable.
First Moment method on $Z$ \text{satisfying assignments of $G_u$}:

$$E_{Z \in 2^n} \{ \frac{2^k}{q} \} \exp \left( \frac{\ln 2}{\alpha \log p_1} \right)$$

The exponent decreases in $\alpha$, crosses zero at $\alpha_1 = 2^k \ln 2 O(p_1 q)$.

First moment threshold $\alpha_1$ separates $E_Z < 8$ and $E_Z < 0$.

Second Moment method: $P_{Z \not\in 0} s \in p E_{Z \in 2^n} q^2 E_{r \in 2^s} s$.

To be useful, requires always $E_{r \in 2^s} s \geq E_{Z \in 2^n} q^2$. Fails, for all $\alpha \not\in 0$.

For random colourings and NAE-SAT, second moment method succeeds up to $\alpha_2 = \alpha_{sat} O(p_1 q)$.
**First Moment method** on $Z \equiv |\{\text{satisfying assignments of } G\}|$:
First Moment method on $Z \equiv |\{\text{satisfying assignments of } G\}|$:

$$\mathbb{E} Z = 2^n(1 - 1/2^k)^m$$
**First Moment method** on $Z \equiv |\{\text{satisfying assignments of } \mathcal{G}\}|$:

$$\mathbb{E}Z = 2^n(1 - 1/2^k)^m = \exp\{n \left[\ln 2 + \alpha \log(1 - 1/2^k)\right]\}$$
**First Moment method** on $Z \equiv |\{\text{satisfying assignments of } G\}|$:

$$E Z = 2^n (1 - 1/2^k)^m = \exp\{n \left[ \ln 2 + \alpha \log(1 - 1/2^k) \right]\}$$

exponent decreases in $\alpha$, crosses zero at $\alpha_1 \approx 2^k \ln 2 + O(1)$
**First Moment method** on $Z \equiv |\{\text{satisfying assignments of } G\}|$:

\[
\mathbb{E}Z = 2^n (1 - 1/2^k)^m = \exp\{n \left[\ln 2 + \alpha \log(1 - 1/2^k)\right]\}
\]

exponent decreases in $\alpha$, crosses zero at $\alpha_1 \approx 2^k \ln 2 + O(1)$

First moment threshold $\alpha_1$ separates $\mathbb{E}Z \to \infty$ $|\mathbb{E}Z \to 0$. 
First Moment method on $Z \equiv |\{\text{satisfying assignments of } \mathcal{G}\}|$:

$$
EZ = 2^n(1 - 1/2^k)^m = \exp\{n\left[\ln 2 + \alpha \log(1 - 1/2^k)\right]\}
$$

exponent decreases in $\alpha$, crosses zero at $\alpha_1 \approx 2^k \ln 2 + O(1)$

First moment threshold $\alpha_1$ separates $EZ \to \infty$ | $EZ \to 0$.

$\alpha_1 \neq \alpha_{sat}$: At least $\epsilon n$ unconstrained variables so $Z > 0 \Rightarrow Z \geq 2^{\epsilon n}$. 
First Moment method on $Z \equiv |\{\text{satisfying assignments of } G\}|$:

$$E[Z] = 2^n(1 - 1/2^k)^m = \exp\{n \left[\ln 2 + \alpha \log(1 - 1/2^k)\right]\}$$

exponent decreases in $\alpha$, crosses zero at $\alpha_1 \approx 2^k \ln 2 + O(1)$

First moment threshold $\alpha_1$ separates $E[Z] \to \infty \mid E[Z] \to 0$.

$\alpha_1 \neq \alpha_{\text{sat}}$: At least $\epsilon n$ unconstrained variables so $Z > 0 \Rightarrow Z \geq 2^{\epsilon n}$.

Second Moment method:

$$\mathbb{P}[Z > 0] \geq \frac{(E[Z])^2}{E[Z^2]}$$
**First Moment method** on $Z \equiv |\{\text{satisfying assignments of } G\}|$:

$$
\mathbb{E} Z = 2^n (1 - 1/2^k)^m = \exp\{n \left[ \ln 2 + \alpha \log(1 - 1/2^k) \right]\}
$$

exponent decreases in $\alpha$, crosses zero at $\alpha_1 \approx 2^k \ln 2 + O(1)$

First moment threshold $\alpha_1$ separates $\mathbb{E} Z \to \infty | \mathbb{E} Z \to 0$.

$\alpha_1 \neq \alpha_{\text{sat}}$: At least $\epsilon n$ unconstrained variables so $Z > 0 \Rightarrow Z \geq 2^{\epsilon n}$.

**Second Moment method**:

$$
\mathbb{P}[Z > 0] \geq \frac{(\mathbb{E} Z)^2}{\mathbb{E}[Z^2]}
$$

To be useful, requires always $\mathbb{E}[Z^2] \geq (\mathbb{E} Z)^2$. Fails, for all $\alpha > 0$. 
**First Moment method** on $Z \equiv |\{\text{satisfying assignments of } G\}|$:

$$\mathbb{E}Z = 2^n(1 - 1/2^k)^m = \exp\left\{n \left[\ln 2 + \alpha \log(1 - 1/2^k)\right]\right\}$$

exponent decreases in $\alpha$, crosses zero at $\alpha_1 \approx 2^k \ln 2 + O(1)$

First moment threshold $\alpha_1$ separates $\mathbb{E}Z \to \infty \big| \mathbb{E}Z \to 0$.

$\alpha_1 \neq \alpha_{\text{sat}}$: At least $\epsilon n$ unconstrained variables so $Z > 0 \Rightarrow Z \geq 2^{\epsilon n}$.

**Second Moment method**: 

$$\mathbb{P}[Z > 0] \geq \frac{(\mathbb{E}Z)^2}{\mathbb{E}[Z^2]}$$

To be useful, requires always $\mathbb{E}[Z^2] \approx (\mathbb{E}Z)^2$. Fails, for all $\alpha > 0$.

For random colourings and NAE-SAT, second moment method succeeds up to $\alpha_2 = \alpha_{\text{sat}} - O(1)$. 
Central complication of random $k$-SAT is that $Z$ fails to concentrate for two distinct (but entangled) reasons —
Central complication of random $k$-SAT is that $Z$ fails to concentrate for two distinct (but entangled) reasons —

1. **condensation and replica symmetry breaking (RSB)** —
   $\mathbb{E}Z$ dominated by **atypical clusters** for $\alpha > \alpha_{\text{sat}} - \epsilon$
Central complication of random $k$-SAT is that $Z$ fails to concentrate for two distinct (but entangled) reasons —

1. **condensation and replica symmetry breaking (RSB)** —
   $\mathbb{E}Z$ dominated by **atypical clusters** for $\alpha > \alpha_{\text{sat}} - \epsilon$

2. **neighborhood profile fluctuations** —
   $\mathbb{E}Z$ dominated by **atypical graphs** for all $\alpha > 0$
Some physics perspective:
condensation and replica symmetry breaking
Spin glasses are marked by a prevalence of frustrated interactions—e.g. Sherrington Kirkpatrick spin-glass ('75): sample $p_{\text{g}}^{ij}$, standard $n_p$, then use them to define $P_{\text{g}}^x q_1 Z \exp \beta (\sum_i \sum_j g_{ij} x_i x_j)$, $p_{t+1}^\pm u_n$.

Some remarkable predictions proved for dense graphs—e.g. for the SK spin-glass Guerra '03, Talagrand '06: Parisi formula (conjecture: Parisi '79, '80) Panchenko '11: Parisi ultrametricity (conjecture: Parisi '79, '80) and for optimization on complete graphs with random edge weights: Aldous '00: random assignment (conjecture: M´ezard–Parisi '85, '86, '87) Frieze '02, W¨astlund '10: TSP (conjecture: M´ezard–Parisi '86, Krauth–M´ezard '89).

Spin glasses are marked by a prevalence of frustrated interactions — e.g. Sherrington Kirkpatrick spin-glass ('75): sample \((g_{ij})_{i<j}\), standard \(N(0, 1)\) then use them to define
Spin glasses are marked by a prevalence of frustrated interactions — e.g. Sherrington Kirkpatrick spin-glass (’75): sample \((g_{ij})_{i<j}\), standard \(N(0,1)\) then use them to define

\[
P(x) \approx \frac{1}{Z} \exp \left\{ \frac{\beta}{\sqrt{n}} \sum_{i<j} g_{ij} x_i x_j \right\}, \quad x \in \{+1, -1\}^n
\]
Spin glasses are marked by a prevalence of frustrated interactions — e.g. Sherrington Kirkpatrick spin-glass ('75): sample \((g_{ij})_{i<j}\), standard \(N(0, 1)\) then use them to define

\[
P(x) \approx \frac{1}{Z} \exp \left\{ \frac{\beta}{\sqrt{n}} \sum_{i<j} g_{ij} x_i x_j \right\}, \quad x \in \{+1, -1\}^n
\]

Some remarkable predictions proved for dense graphs — e.g. for the SK spin-glass

Guerra '03, Talagrand '06: Parisi formula (conjecture: Parisi '79, '80)
Panchenko '11: Parisi ultrametricity (conjecture: Parisi '79, '80)

and for optimization on complete graphs with random edge weights:

Aldous '00: random assignment (conjecture: Mézard–Parisi '85, '86, '87)
Frieze '02, Wästlund '10: TSP (conjecture: Mézard–Parisi '86, Krauth–Mézard '89)
Spin glasses are marked by a prevalence of frustrated interactions — e.g. Sherrington Kirkpatrick spin-glass (‘75): sample $(g_{ij})_{i<j}$, standard $N(0,1)$ then use them to define

$$\mathbb{P}(x) \equiv \frac{1}{Z} \exp\left\{ \frac{\beta}{\sqrt{n}} \sum_{i<j} g_{ij} x_i x_j \right\}, \quad x \in \{+1, -1\}^n$$

Some remarkable predictions proved for dense graphs — e.g. for the SK spin-glass

Guerra ‘03, Talagrand ‘06: Parisi formula (conjecture: Parisi ‘79, ‘80)
Panchenko ‘11: Parisi ultrametricity (conjecture: Parisi ‘79, ‘80)

and for optimization on complete graphs with random edge weights:

Aldous ‘00: random assignment (conjecture: Mézard–Parisi ‘85, ‘86, ‘87)
Frieze ‘02, Wästlund ‘10: TSP (conjecture: Mézard–Parisi ‘86, Krauth–Mézard ‘89)

More recently a set of predictions for sparse random systems emerged:

Krząkała–Montanari–Ricci-Tersenghi–Semerjian–Zdeborová ‘07,
Montanari–Ricci-Tersenghi–Semerjian ‘08
Phase Diagram

Two solutions are connected if they differ by one bit.

KMRSZ '07, MRS '08
Phase Diagram

The solution space SOL starts out as a well-connected cluster.
Phase Diagram

The solution space **SOL** starts out as a well-connected cluster.

After $\alpha_{\text{clust}}$, **SOL** decomposes into exponentially clusters

–Clustering Achlioptas, Coja-Oghlan '10
The solution space SOL starts out as a well-connected cluster.

After $\alpha_{\text{clust}}$, SOL decomposes into exponentially clusters

$\cdots$Clustering Achlioptas, Coja-Oghlan '10

After $\alpha_{\text{cond}}$, SOL is dominated by a few large clusters
Phase Diagram

The solution space **SOL** starts out as a well-connected cluster.

After $\alpha_{\text{clust}}$, **SOL** decomposes into exponentially clusters

– Clustering Achlioptas, Coja-Oghlan '10

After $\alpha_{\text{cond}}$, **SOL** is dominated by a few large clusters

After $\alpha_{\text{sat}}$, no solutions w.h.p.
**Condensation** (in regular models)

Complexity function $\Sigma \equiv \Sigma_\alpha(s)$ such that:

$$
\mathbb{E}Z = \sum_{\text{cluster size}} \exp\{ns\} \times \mathbb{E}[\text{number of clusters of that size}] \times \exp\{n\Sigma(s)\}
$$
**Condensation** (in regular models)

Complexity function $\Sigma \equiv \Sigma_\alpha(s)$ such that:

$$\mathbb{E} Z = \sum \underbrace{(\text{cluster size}) \times \mathbb{E}[\text{number of clusters of that size}]}_{\exp\{ns\}} \underbrace{\exp\{n\Sigma(s)\}}_{\exp\{n\Sigma(s)\}}$$
Condensation (in regular models)
Complexity function $\Sigma \equiv \Sigma_\alpha(s)$ such that:

$$
\mathbb{E}Z = \sum \text{(cluster size)} \times \mathbb{E}[\text{number of clusters of that size}] 
\exp\{ns\} \times \exp\{n\Sigma(s)\}
$$

$\mathbb{E}Z$ is dominated by $s$ where $\Sigma'(s) \equiv -1$ (depending on $\alpha$).
**Condensation** (in regular models)

Complexity function $\Sigma \equiv \Sigma_\alpha(s)$ such that:

$$\mathbb{E}Z = \sum (\text{cluster size}) \times \mathbb{E}[^\text{number of clusters of that size}]$$

$$\exp\{ns\} \times \exp\{n\Sigma(s)\}$$

$\mathbb{E}Z$ is dominated by $s$ where $\Sigma'(s) \equiv -1$ (depending on $\alpha$).
Condensation (in regular models)

Complexity function $\Sigma \equiv \Sigma_\alpha(s)$ such that:

$$\mathbb{E} Z = \sum \text{(cluster size)} \times \mathbb{E} \left[ \text{number of clusters of that size} \right] \text{exp}\{ns\} \text{exp}\{n\Sigma(s)\}$$

$\mathbb{E} Z$ is dominated by $s$ where $\Sigma'(s) \equiv -1$ (depending on $\alpha$).

![Diagram of condensation](image)
**Condensation** (in regular models)

Complexity function $\Sigma \equiv \Sigma_\alpha(s)$ such that:

$$
\mathbb{E} Z = \sum (\text{cluster size}) \times \mathbb{E} [\text{number of clusters of that size}] = \sum \exp\{ns\} \times \exp\{n\Sigma(s)\}
$$

$\mathbb{E} Z$ is dominated by $s$ where $\Sigma'(s) \equiv -1$ (depending on $\alpha$).
Condensation (in regular models)

Complexity function $\Sigma \equiv \Sigma_\alpha(s)$ such that:

$$\mathbb{E}Z = \sum \text{(cluster size)} \times \mathbb{E}[\text{number of clusters of that size}] \exp\{ns\} \exp\{n\Sigma(s)\}$$

$\mathbb{E}Z$ is dominated by $s$ where $\Sigma'(s) \equiv -1$ (depending on $\alpha$).
Condensation (in regular models)

Complexity function $\Sigma \equiv \Sigma_\alpha(s)$ such that:

$$E Z = \sum (\text{cluster size}) \times E[\text{number of clusters of that size}] \exp\{ns\} \exp\{n\Sigma(s)\}$$

$E Z$ is dominated by $s$ where $\Sigma'(s) \equiv -1$ (depending on $\alpha$).
Condensation (in regular models)

Complexity function $\Sigma \equiv \Sigma_\alpha(s)$ such that:

$$\mathbb{E}Z = \sum \text{(cluster size)} \times \mathbb{E}[\text{number of clusters of that size}]$$

$$\exp\{n \Sigma(s)\}$$

$\mathbb{E}Z$ is dominated by $s$ where $\Sigma'(s) \equiv -1$ (depending on $\alpha$).
**Condensation** (in regular models)

Complexity function $\Sigma \equiv \Sigma_\alpha(s)$ such that:

$$
\mathbb{E}Z = \sum_{\text{(cluster size)}} \times \mathbb{E}[\text{number of clusters of that size}] \\
\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quadende
Condensation and non-concentration

The 1-RSB prediction:

- Satisfiability Threshold
  \[ \alpha_{\text{sat}} = \sup \{ \alpha : \sup_s \Sigma(s) \geq 0 \} \]

- Condensation Threshold and free energy
  \[ \alpha_{\text{cond}} = \sup \{ \alpha : \sup_s s + \Sigma(s) = \sup_{s: \Sigma(s) \geq 0} s + \Sigma(s) \} \]
  \[ \Phi = \lim_{n \to \infty} \frac{1}{n} \log Z = \sup \{ s + \Sigma(s) : \Sigma(s) > 0 \} = \sup \{ s : \Sigma(s) > 0 \} \]
RSB: 

The one step replica symmetry breaking (1RSB) heuristic roughly says there is no extra structure at the cluster level and decay of correlation. The 1RSB heuristic suggests that, absent other sources (neighborhood fluctuations) of non-concentration, we can locate exact threshold by moment method on \( \Omega \) #clusters.

Results beyond the condensation threshold:
- Condensation Threshold: Random \( k \)-Colourings \( G(n,p) \) large \( k \) [Bapst, Coja-Oghlan, Hetterich, Rassmann, Vilenchik]
- Condensation Regime Free Energy: Regular \( k \)-NAESAT large \( k \) [S’, Sun, Zhang]
- Satisfiability Threshold: Regular NAESAT large \( k \) [Ding, S’, Sun]
- Maximum Independent Set \( d \)-Regular, large \( d \) [Ding, S’, Sun]
- Regular SAT, large \( k \) [Coja-Oghlan, Panagiotou]
- Random \( k \)-SAT, large \( k \) [Ding, S’, Sun]
RSB: The one step replica symmetry breaking (1RSB) heuristic roughly says there is no extra structure at the cluster level and decay of correlation.
RSB: The one step replica symmetry breaking (1RSB) heuristic roughly says there is no extra structure at the cluster level and decay of correlation.

The 1RSB heuristic suggests that, absent other sources (neighborhood fluctuations) of non-concentration, we can locate exact threshold by moment method on $\Omega \equiv \#\text{clusters}$.
RSB: The one step replica symmetry breaking (1RSB) heuristic roughly says there is no extra structure at the cluster level and decay of correlation.

The 1RSB heuristic suggests that, absent other sources (neighborhood fluctuations) of non-concentration, we can locate exact threshold by moment method on $\Omega \equiv \#\text{clusters}$.

**Results beyond the condensation threshold:**

Condensation Threshold:

- Random $k$-Colourings $G(n,p)$ large $k$ 
  [Bapst, Coja-Oghlan, Hetterich, Rassmann, Vilenchik]
- Regular $k$-NAESAT large $k$ 
  [S’, Sun, Zhang]
RSB: The one step replica symmetry breaking (1RSB) heuristic roughly says there is no extra structure at the cluster level and decay of correlation.

The 1RSB heuristic suggests that, absent other sources (neighborhood fluctuations) of non-concentration, we can locate exact threshold by moment method on $\Omega \equiv \#\text{clusters}$.

**Results beyond the condensation threshold:**

Condensation Threshold:
- Random $k$-Colourings $G(n,p)$ large $k$
  - [Bapst, Coja-Oghlan, Hetterich, Rassmann, Vilenchik]
- Regular $k$-NAESAT large $k$
  - [S’, Sun, Zhang]

Condensation Regime Free Energy:
- Regular $k$-NAESAT large $k$
  - [S’, Sun, Zhang]
**RSB:** The one step replica symmetry breaking (1RSB) heuristic roughly says there is no extra structure at the cluster level and decay of correlation.

The 1RSB heuristic suggests that, *absent other sources (neighborhood fluctuations) of non-concentration*, we can locate exact threshold by moment method on $\Omega \equiv \#\text{clusters}$.

**Results beyond the condensation threshold:**

Condensation Threshold:
- Random $k$-Colourings $G(n,p)$ large $k$
  - [Bapst, Coja-Oghlan, Hetterich, Rassmann, Vilenchik]
- Regular $k$-NAESAT large $k$
  - [S’, Sun, Zhang]

Condensation Regime Free Energy:
- Regular $k$-NAESAT large $k$
  - [S’, Sun, Zhang]

Satisfiability Threshold:
- Regular NAESAT large $k$
  - [Ding, S’, Sun]
- Maximum Independent Set $d$-Regular, large $d$
  - [Ding, S’, Sun]
- Regular SAT, large $k$
  - [Coja-Oghlan, Panagiotou]
- Random $k$-SAT, large $k$
  - [Ding, S’, Sun]
**Encoding of local neighborhood** We represent clusters as a new spin system on $V(G)$. 

Start from $x_P \pm u V_p G_q$ and explore the cluster $C$. If a spin can be flipped between $\pm$ without violating any clauses it is set to $f$. Iterate until done. Each variable is mapped to a value from $t, -u, f$. This resulting configuration on $t, -u, f V_p G_q$ is our definition of a cluster. It is a spin system satisfying the following conditions: $f$ are not forced by any clause. $\pm$ and $-u$ variables must be forced by at least one clause. No violated clause. Locally rigid resulting in no clustering.
Encoding of local neighborhood  We represent clusters as a new spin system on $V(G)$.

- Start from $x \in \{+, -\}^{V(G)}$ and explore the cluster $C$. 
**Encoding of local neighborhood** We represent clusters as a new spin system on $V(G)$.

- Start from $x \in \{+, -\}^{V(G)}$ and explore the cluster $C$.
- If a spin can be flipped between $+$ and $-$ without violating any clauses it is set to $f$.
- Iterate until done.
Encoding of local neighborhood We represent clusters as a new spin system on \( V(G) \).

- Start from \( x \in \{+, -\}^{V(G)} \) and explore the cluster \( C \).
- If a spin can be flipped between + and - without violating any clauses it is set to \( f \).
- Iterate until done.
- Each variable is mapped to a value from \( \{+, -, f\} \).
**Encoding of local neighborhood** We represent clusters as a new spin system on $V(G)$.

- Start from $x \in \{+,-\}^{V(G)}$ and explore the cluster $C$.
- If a spin can be flipped between $+$ and $-$ without violating any clauses it is set to $f$.
- Iterate until done.
- Each variable is mapped to a value from $\{+,-,f\}$.

This resulting configuration on $\{+,-,f\}^{V(G)}$ is our definition of a cluster. It is a spin system satisfying the following conditions:

- $f$ are not forced by any clause.
- $+$ and $-$ variables must be forced by at least one clause.
- No violated clause.
**Encoding of local neighborhood** We represent clusters as a new spin system on $V(G)$.

- Start from $x \in \{+, -\}^{V(G)}$ and explore the cluster $C$.
- If a spin can be flipped between $+$ and $-$ without violating any clauses it is set to $f$.
- Iterate until done.
- Each variable is mapped to a value from $\{+, -, f\}$.

This resulting configuration on $\{+, -, f\}^{V(G)}$ is our definition of a cluster. It is a spin system satisfying the following conditions:

- $f$ are not forced by any clause.
- $+$ and $-$ variables must be forced by at least one clause.
- No violated clause.

Locally rigid resulting in no clustering.
**Cavity Method**: Adding a new vertex \( v \) (or clause).
**Cavity Method**: Adding a new vertex $v$ (or clause).
**Cavity Method**: Adding a new vertex \( v \) (or clause).

If we know the joint distribution of \( \sigma_{u_i} \) we can:

1. Calculate the law of \( \sigma_v \)
2. Evaluate the change in the partition function from \( Z_{n+1}/Z_n \).

Write \( \log Z_n = \sum_{i=1}^{n} \log Z_i/Z_{i-1} \).
**Cavity Method**: Adding a new vertex $v$ (or clause).

If we know the joint distribution of $\sigma_{u_i}$ we can:

1. Calculate the law of $\sigma_v$
2. Evaluate the change in the partition function from $Z_{n+1}/Z_n$.

Write $\log Z_n = \sum_{i=1}^{n} \log Z_i/Z_{i-1}$.

The **Replica Symmetric** heuristic assumes that $\sigma_{u_i}$ are independent drawn from some law $\mu$. 

\[ \text{Diagram showing added vertex } v \text{ and connected vertices } u_1, u_2, u_3, u_4. \]
**Cavity Method**: Adding a new vertex $v$ (or clause).

If we know the joint distribution of $\sigma_{u_i}$ we can:

1. Calculate the law of $\sigma_v$
2. Evaluate the change in the partition function from $Z_{n+1}/Z_n$.

Write $\log Z_n = \sum_{i=1}^n \log Z_i/Z_{i-1}$.

The **Replica Symmetric** heuristic assumes that $\sigma_{u_i}$ are independent drawn from some law $\mu$.

The **1-RSB** heuristic assumes this for the cluster model.
**Cavity Method**: Adding a new vertex $v$ (or clause).

If we know the joint distribution of $\sigma_{u_i}$ we can:

1. Calculate the law of $\sigma_v$
2. Evaluate the change in the partition function from $Z_{n+1}/Z_n$.

Write $\log Z_n = \sum_{i=1}^{n} \log Z_i/Z_{i-1}$.

The **Replica Symmetric** heuristic assumes that $\sigma_{u_i}$ are independent drawn from some law $\mu$.

The **1-RSB** heuristic assumes this for the cluster model.

**Self-consistency**: The law of $\sigma_v$ should also be drawn from $\mu$ which means $\mu$ must satisfy a fixed point equation.
Explicit formula \((k \geq 3)\)
**Explicit formula \((k \geq 3)\)** Let \(\mathcal{P} \equiv \) space of probability measures on \([0, 1]\). Define the distributional recursion \(R_\alpha : \mathcal{P} \to \mathcal{P}\),

\[
R_\alpha \mu(B) \equiv \sum_{d \equiv (d^+, d^-)} \pi_\alpha(d) \int 1 \left\{ \frac{(1 - \Pi^-)\Pi^+}{\Pi^+ + \Pi^- - \Pi^+\Pi^-} \in B \right\} \prod_{i,j} d\mu(\eta^\pm_{ij})
\]

with \(\pi_\alpha(d) \equiv \frac{e^{-k\alpha(k\alpha/2)d^+d^-}}{(d^+)! (d^-)!}\), \(\Pi^\pm \equiv \Pi^\pm(d, \eta) \equiv \prod_{i=1}^{d^*} \left(1 - \prod_{j=1}^{k-1} \eta^\pm_{ij}\right)\)

We show \((R_\alpha)^\ell 1_{1/2} \xrightarrow{\ell \to \infty} \mu_\alpha\).

*Distributional equation for the chance of being + in a random cluster.*
Explicit formula \((k \geq 3)\) Let \(\mathcal{P} \equiv \) space of probability measures on \([0, 1]\). Define the distributional recursion \(R_\alpha : \mathcal{P} \rightarrow \mathcal{P}\),

\[
R_\alpha \mu(B) \equiv \sum_{d \equiv (d^+, d^-)} \pi_\alpha(d) \int 1 \left\{ \frac{(1 - \Pi^-) \Pi^+}{\Pi^+ + \Pi^- - \Pi^+ \Pi^-} \in B \right\} \prod_{i,j} d\mu(\eta_{ij}^\pm)
\]

with \(\pi_\alpha(d) \equiv \frac{e^{-k\alpha(k\alpha/2)d^+d^-}}{(d^+)! (d^-)!}\), \(\Pi^\pm \equiv \Pi^\pm(d, \eta) \equiv \prod_{i=1}^{d^*} \left( 1 - \prod_{j=1}^{k-1} \eta_{ij}^\pm \right)\)

We show \((R_\alpha)^\ell 1_{1/2} \overset{\ell \to \infty}{\longrightarrow} \mu_\alpha\).

Distributional equation for the chance of being \(+\) in a random cluster.

Define

\[
\Phi(\alpha) = \sum_d \pi_\alpha(d) \int \ln \left( \Pi^+ + \Pi^- - \Pi^+ \Pi^- \right) \prod_j d\mu_\alpha(\eta_j) \prod_{i,j} d\mu_\alpha(\eta_{ij}^\pm) \]

\[-\alpha(k - 1) \int \ln \left( 1 - \prod_{j=1}^{k} \eta_j \right) \prod_j d\mu_\alpha(\eta_j) \prod_{i,j} d\mu_\alpha(\eta_{ij}^\pm)\]

Expected change in \(\log \Omega_n\) to \(\log \Omega_{n+1}\).
Explicit formula \((k \geq 3)\) Let \(\mathcal{P} \equiv \) space of probability measures on \([0, 1]\). Define the distributional recursion \(R_{\alpha} : \mathcal{P} \to \mathcal{P}\),

\[
R_{\alpha}\mu(B) \equiv \sum_{d \equiv (d^+, d^-)} \pi_\alpha(d) \int 1 \left\{ \frac{(1 - \Pi^-)\Pi^+}{\Pi^+ + \Pi^- - \Pi^+\Pi^-} \in B \right\} \prod_{i,j} d\mu(\eta_{ij}^*)
\]

with \(\pi_\alpha(d) \equiv \frac{e^{-k\alpha(k\alpha/2)d^+d^-}}{(d^+)! (d^-)!}\), \(\Pi^+ \equiv \Pi^+(d, \eta) \equiv \prod_{i=1}^{d^*} \left( 1 - \prod_{j=1}^{k-1} \eta_{ij}^* \right)\)

We show \(\left(R_{\alpha}\right)^{\ell} 1_{1/2} \xrightarrow{\ell \to \infty} \mu_\alpha\).

Distributional equation for the chance of being \(+\) in a random cluster.

Define

\[
\Phi(\alpha) = \sum_d \pi_\alpha(d) \int \ln \left( \Pi^+ + \Pi^- - \Pi^+\Pi^- \right) \prod_j d\mu_\alpha(\eta_j) \prod_{i,j} d\mu_\alpha(\eta_{ij}^*)
\]

\[-\alpha(k - 1) \int \ln \left( 1 - \prod_{j=1}^{k} \eta_j \right) \prod_j d\mu_\alpha(\eta_j) \prod_{i,j} d\mu_\alpha(\eta_{ij}^*)\]

Expected change in \(\log \Omega_n\) to \(\log \Omega_{n+1}\).

Then the 1RSB prediction \(\alpha_{\text{sat}} \approx 2^k \ln 2 - (1 + \ln 2)/2\) is the root of \(\Phi(\alpha) = 0\).
Previous Bounds: Satisfiability conjecture is known in special case $k = 2$, with $\alpha_{\text{sat}} = 1$

Goerdt '92, '96, Chvátal–Reed '92, de la Vega '92
**Previous Bounds:** Satisfiability conjecture is known in special case $k = 2$, with $\alpha_{\text{sat}} = 1$

Goerdt '92, '96, Chvátal–Reed '92, de la Vega '92

Bounds for $k \geq 3$: $(\epsilon_k \to 0$ as $k \to \infty)$
**Previous Bounds:** Satisfiability conjecture is known in special case $k = 2$, with $\alpha_{sat} = 1$

Goerdt '92, '96, Chvátal–Reed '92, de la Vega '92

Bounds for $k \geq 3$: $(\varepsilon_k \to 0$ as $k \to \infty)$

<table>
<thead>
<tr>
<th>$\alpha_{sat}$</th>
<th>bound on threshold</th>
<th>gap</th>
<th>trivial</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\leq$</td>
<td>$2^k \ln 2 - (\ln 2)/2 + \varepsilon_k$</td>
<td>$O(1)$</td>
<td></td>
</tr>
</tbody>
</table>
Previous Bounds: Satisfiability conjecture is known in special case $k = 2$, with $\alpha_{\text{sat}} = 1$

Goerdt '92, '96, Chvátal–Reed '92, de la Vega '92

Bounds for $k \geq 3$: ($\epsilon_k \to 0$ as $k \to \infty$)

<table>
<thead>
<tr>
<th>$\alpha_{\text{sat}} \leq$</th>
<th>bound on threshold</th>
<th>gap</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$2^k \ln 2 - (\ln 2)/2 + \epsilon_k$</td>
<td>$O(1)$</td>
</tr>
<tr>
<td></td>
<td>$2^k \ln 2 - (1 + \ln 2)/2 + \epsilon_k$</td>
<td>$\epsilon_k$</td>
</tr>
</tbody>
</table>
### Previous Bounds: Satisfiability conjecture is known in special case \( k = 2 \), with \( \alpha_{\text{sat}} = 1 \)

Goerdt ’92, ’96, Chvátal–Reed ’92, de la Vega ’92

#### Bounds for \( k \geq 3 \): \( (\epsilon_k \rightarrow 0 \text{ as } k \rightarrow \infty) \)

<table>
<thead>
<tr>
<th>( \alpha_{\text{sat}} \leq )</th>
<th>bound on threshold</th>
<th>gap</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 2^k \ln 2 - (\ln 2)/2 + \epsilon_k )</td>
<td>( O(1) )</td>
<td>trivial</td>
</tr>
<tr>
<td>( 2^k \ln 2 - (1 + \ln 2)/2 + \epsilon_k )</td>
<td>( \epsilon_k )</td>
<td>Kirousis et al. ’98</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \alpha_{\text{sat}} \geq )</th>
<th>(algorithmic) ( 1.817 \cdot 2^k/k )</th>
<th>( 2^k \ln 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(algorithmic) ( 2^k (\ln k)/k )</td>
<td>( 2^k \ln 2 )</td>
<td>Frieze–Suen ’96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coja-Oghlan ’10</td>
</tr>
</tbody>
</table>
**Previous Bounds:** Satisfiability conjecture is known in special case $k = 2$, with $\alpha_{\text{sat}} = 1$

Goerdt '92, '96, Chvátal–Reed '92, de la Vega '92

Bounds for $k \geq 3$:  

$$\left(\epsilon_k \rightarrow 0 \text{ as } k \rightarrow \infty\right)$$

| $\alpha_{\text{sat}} \leq$ | bound on threshold | gap | \quad \quad \quad \quad | \\
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$2^k \ln 2 - (\ln 2)/2 + \epsilon_k$</td>
<td>$O(1)$</td>
<td>$\epsilon_k$</td>
<td>trivial $\quad \quad \quad \quad$ Kirousis et al. '98</td>
</tr>
<tr>
<td>$2^k \ln 2 - (1 + \ln 2)/2 + \epsilon_k$</td>
<td>$O(1)$</td>
<td>$\epsilon_k$</td>
<td>trivial $\quad \quad \quad \quad$ Kirousis et al. '98</td>
</tr>
</tbody>
</table>

| $\alpha_{\text{sat}} \geq$ | bound on threshold | gap | \quad \quad \quad \quad | \\
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(algorithmic) $1.817 \cdot 2^k/k$</td>
<td>$2^k \ln 2$</td>
<td>$2^k \ln 2$</td>
<td>Frieze–Suen '96 $\quad \quad \quad \quad$ Coja-Oghlan '10</td>
</tr>
<tr>
<td>(algorithmic) $2^k (\ln k)/k$</td>
<td>$2^k \ln 2$</td>
<td>$2^k \ln 2$</td>
<td>Frieze–Suen '96 $\quad \quad \quad \quad$ Coja-Oghlan '10</td>
</tr>
<tr>
<td>$2^{k-1} \ln 2 - O(1)$</td>
<td>$2^k \ln 2$</td>
<td>$2^k \ln 2$</td>
<td>Frieze–Suen '96 $\quad \quad \quad \quad$ Coja-Oghlan '10</td>
</tr>
<tr>
<td>$2^k \ln 2 - O(k)$</td>
<td>$O(k)$</td>
<td>$O(k)$</td>
<td>Achlioptas–Moore '02 $\quad \quad \quad \quad$ Achlioptas–Peres '03</td>
</tr>
<tr>
<td>$2^k \ln 2 - 3(\ln 2)/2 - \epsilon_k$</td>
<td>$O(1)$</td>
<td>$O(1)$</td>
<td>Achlioptas–Moore '02 $\quad \quad \quad \quad$ Achlioptas–Peres '03</td>
</tr>
<tr>
<td>$2^k \ln 2 - (1 + \ln 2)/2 - \epsilon_k$</td>
<td>$\epsilon_k$</td>
<td>$\epsilon_k$</td>
<td>Achlioptas–Moore '02 $\quad \quad \quad \quad$ Achlioptas–Peres '03</td>
</tr>
</tbody>
</table>

---

RSB: Satisfiability bounds (18/20)
**Previous Bounds:** Satisfiability conjecture is known in special case \( k = 2 \), with \( \alpha_{\text{sat}} = 1 \)

Goerdt '92, '96, Chvátal–Reed '92, de la Vega '92

Bounds for \( k \geq 3 \):

\[
(\epsilon_k \rightarrow 0 \text{ as } k \rightarrow \infty)
\]

<table>
<thead>
<tr>
<th>( \alpha_{\text{sat}} \leq )</th>
<th>( 2^k \ln 2 - (\ln 2)/2 + \epsilon_k )</th>
<th>( O(1) )</th>
<th>trivial</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_{\text{sat}} \geq )</td>
<td>( 2^k \ln 2 - (1 + \ln 2)/2 + \epsilon_k )</td>
<td>( \epsilon_k )</td>
<td>Kirousis et al. '98</td>
</tr>
<tr>
<td>(algorithmic) ( 1.817 \cdot 2^k/k )</td>
<td>( 2^k \ln 2 )</td>
<td>Frieze–Suen '96</td>
<td></td>
</tr>
<tr>
<td>(algorithmic) ( 2^k (\ln k)/k )</td>
<td>( 2^k \ln 2 )</td>
<td>Coja-Oghlan '10</td>
<td></td>
</tr>
<tr>
<td>( 2^{k-1} \ln 2 - O(1) )</td>
<td>( 2^{k-1} \ln 2 )</td>
<td>Achlioptas–Moore '02</td>
<td></td>
</tr>
<tr>
<td>( 2^k \ln 2 - O(k) )</td>
<td>( O(k) )</td>
<td>Achlioptas–Peres '03</td>
<td></td>
</tr>
<tr>
<td>( 2^k \ln 2 - (3\ln 2)/2 - \epsilon_k )</td>
<td>( O(1) )</td>
<td>Coja-Oghlan–Panagiotou '13, '14</td>
<td></td>
</tr>
<tr>
<td>( 2^k \ln 2 - (1 + \ln 2)/2 - \epsilon_k )</td>
<td>( \epsilon_k )</td>
<td>exact threshold</td>
<td></td>
</tr>
<tr>
<td>( \alpha_{\text{sat}} = )</td>
<td>( \alpha_*(k \geq k_0) )</td>
<td>( 0 ) ((k \geq k_0))</td>
<td></td>
</tr>
</tbody>
</table>
Theorem. (Ding, S., Sun) For \( k \geq k_0 \) (absolute constant), random \( k \)-SAT has a sharp satisfiability threshold, with explicit value \( \alpha_{\text{sat}} = \alpha_* \) matching the one-step replica symmetry breaking prediction of Mertens–Mézard–Zecchina ’06.
Open problems:
Open problems:

Other models, random graph chromatic number?
Open problems:

Other models, random graph chromatic number?

Other aspects of the 1RSB phase diagram, condensation to a finite number of clusters?

\[
\frac{1}{N} \sum_{x} \log x^{\alpha} \text{ with } \alpha = \frac{1}{\log(1/2)}.
\]
Open problems:

Other models, random graph chromatic number?

Other aspects of the 1RSB phase diagram, condensation to a finite number of clusters?

Models at finite temperature? Clusters no longer have rigid combinatorial description.
Open problems:

Other models, random graph chromatic number?

Other aspects of the 1RSB phase diagram, condensation to a finite number of clusters?

Models at finite temperature? Clusters no longer have rigid combinatorial description.

Models with full Replica Symmetry breaking, e.g. MAX-CUT?
Thanks!