Clusters of solutions to random linear equations

Michael Molloy

Dept of Computer Science
University of Toronto

includes work with Dimitris Achlioptas and with Jane Gao

A Random System of Linear Equations

We have $M = cn$ linear equations over n $\{0, 1\}$ variables.

All addition is $\mod 2$.

$$\begin{aligned} x_5 + x_1 + x_6 &= 0 \\ x_2 + x_6 + x_1 &= 1 \\ x_1 + x_2 + x_5 &= 1 \\ x_3 + x_5 + x_4 &= 0 \\ x_6 + x_4 + x_2 &= 1 \end{aligned}$$

Each equation contains

- a random k-tuple of variables
- a random $\{0, 1\}$ RHS.

A Random System of Linear Equations

We have $M = cn$ linear equations over n $\{0, 1\}$ variables.

All addition is $\mod 2$.

$$\begin{aligned} x_5 + x_1 + x_6 &= 0 \\ x_2 + x_6 + x_1 &= 1 \\ x_1 + x_2 + x_5 &= 1 \\ x_3 + x_5 + x_4 &= 0 \\ x_6 + x_4 + x_2 &= 1 \end{aligned}$$

Each equation contains

- a random k-tuple of variables
- a random $\{0, 1\}$ RHS.

Also known as k-XORSAT.

Random Constraint Satisfaction Problems

This is one of the standard models of random constraint satisfaction problems.

It is one of the simplest of the commonly studied models:

- lots of symmetry
- solutions are well-understood

Random Constraint Satisfaction Problems

This is one of the standard models of random constraint satisfaction problems.

It is one of the simplest of the commonly studied models:

- lots of symmetry
- solutions are well-understood

We've been able to prove things here that we can't yet prove for, eg., random k-SAT and random graph colouring.

Random Constraint Satisfaction Problems

This is one of the standard models of random constraint satisfaction problems.

It is one of the simplest of the commonly studied models:

- lots of symmetry
- solutions are well-understood

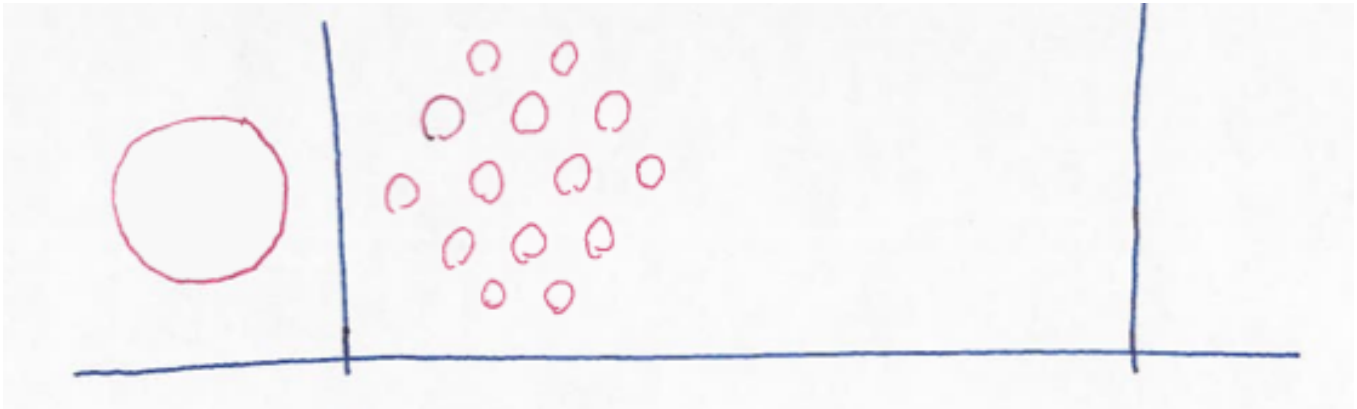
We've been able to prove things here that we can't yet prove for, eg., random k-SAT and random graph colouring.

Satisfiability Threshold:

$$c = .917..., \qquad k = 3 \qquad \text{Dubois and Mandler, 2002}$$

$k > 3$ Dietzfelbinger, et al, 2010, Pittel and Sorkin, 2012.

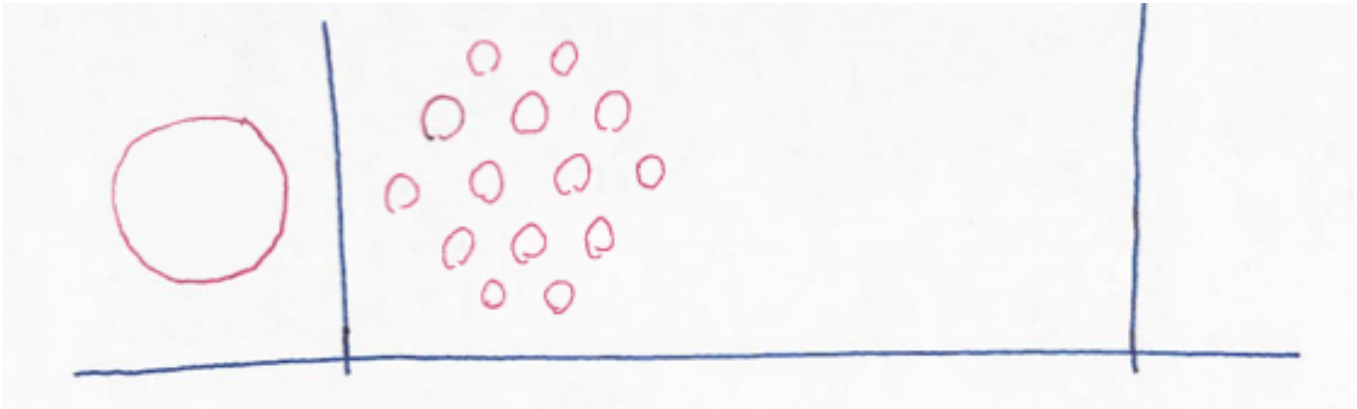
Clustering



c_c c_s

Phenomenon seems to hold for a wide variety of random CSP's.

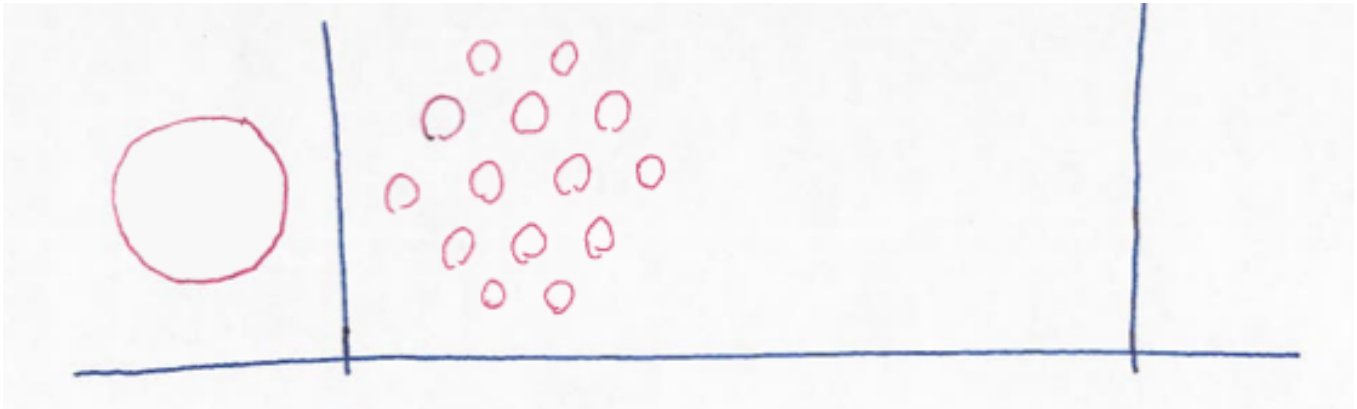
Clustering



c_c c_s

- Well-connected. One can move throughout the cluster changing $o(n)$ variables at a time.
- Well-separated Moving from one cluster to another requires changing $\Theta(n)$ variables in one step.

Clustering

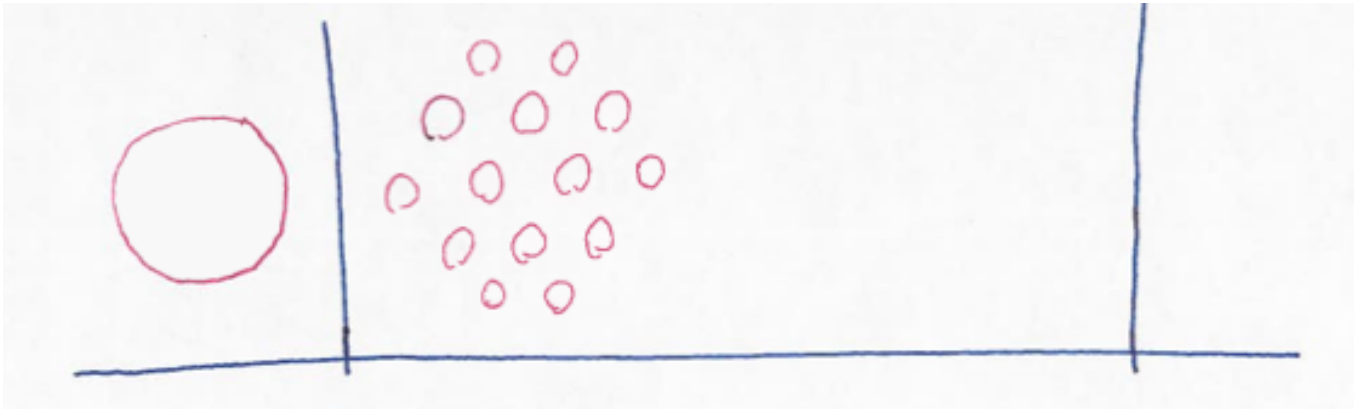


c_c c_s

This is mostly non-rigorous, but:

- It is based on some substantial mathematical analysis
- It explains a lot
  - earlier results
  - algorithmic challenges
- "Knowing" that it is true can inspire proof approaches (eg. the previous talk)
- Understanding random CSP's will require understanding clustering

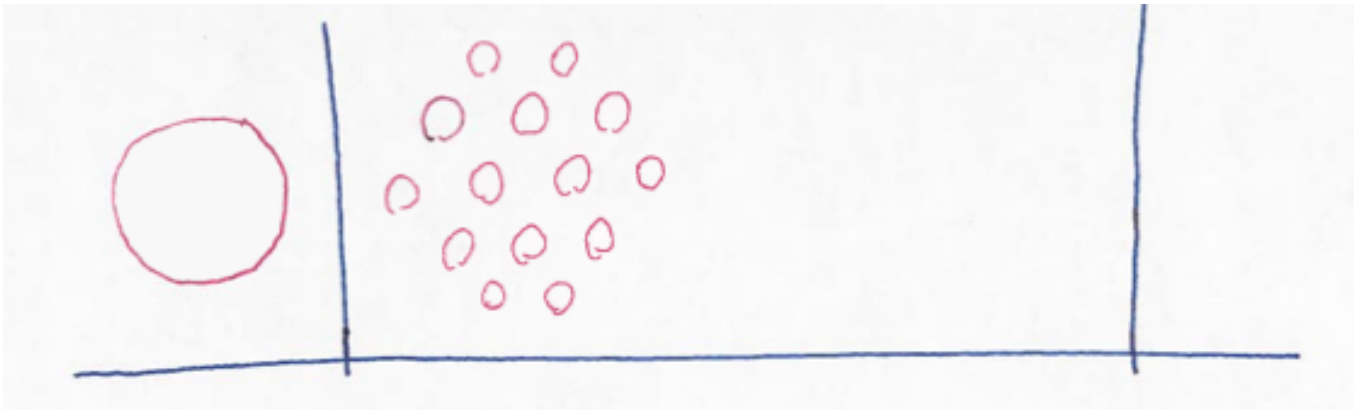
Clustering



c_c c_s

- Well-connected. One can move throughout the cluster changing $o(n)$ variables at a time.
- Well-separated Moving from one cluster to another requires changing $\Theta(n)$ variables in one step.

Clustering



c_c c_s

- Well-connected. One can move throughout the cluster changing $O(\log n)$ variables at a time.
- Well-separated Moving from one cluster to another requires changing $\Theta(n)$ variables in one step.

Ibrahimi, Kanoria, Kranning and Montanari (2011)
Achlioptas and M (2011)

The 2-core

Remove every variable that appears in at most one equation, along with the equation it belongs to.

$$
\begin{aligned}
x_5 + x_1 + x_6 &= 0 \\
x_2 + x_6 + x_1 &= 1 \\
x_1 + x_2 + x_5 &= 1 \\
x_3 + x_5 + x_4 &= 0 \qquad \leftarrow \text{Remove} \\
x_4 + x_6 + x_2 &= 1
\end{aligned}
$$

Reason: Every solution to what remains can easily be extended to the original system, by setting the deleted variable.

The 2-core

Remove every variable that appears in at most one equation, along with the equation it belongs to.

$$
\begin{aligned}
x_5 + x_1 + x_6 &= 0 \\
x_2 + x_6 + x_1 &= 1 \\
x_1 + x_2 + x_5 &= 1 \\
x_3 + x_5 + x_4 &= 0 \qquad \leftarrow \text{Remove} \\
x_6 + x_4 + x_2 &= 1 \qquad \leftarrow \text{Remove}
\end{aligned}
$$

Iterate

The 2-core

Remove every variable that appears in at most one equation, along with the equation it belongs to.

$$
\begin{aligned}
x_5 + x_1 + x_6 &= 0 \\
x_2 + x_6 + x_1 &= 1 \\
x_1 + x_2 + x_5 &= 1 \\
x_3 + x_5 + x_4 &= 0 \qquad \leftarrow \text{Remove} \\
x_6 + x_4 + x_2 &= 1 \qquad \leftarrow \text{Remove}
\end{aligned}
$$

What remains is the 2-core of the system.

The 2-core

Remove every variable that appears in at most one equation, along with the equation it belongs to.

$$
\begin{aligned}
x_5 + x_1 + x_6 &= 0 \\
x_2 + x_6 + x_1 &= 1 \\
x_1 + x_2 + x_5 &= 1 \\
x_3 + x_5 + x_4 &= 0 \qquad \leftarrow \text{Remove} \\
x_6 + x_4 + x_2 &= 1 \qquad \leftarrow \text{Remove}
\end{aligned}
$$

What remains is the 2-core of the system.

This is also the 2-core of the underlying hypergraph:

- vertices are the variables
- hyperedges are the k-tuples of vertices that form equations

The 2-core

Remove every variable that appears in at most one equation, along with the equation it belongs to.

$$
\begin{aligned}
x_5 + x_1 + x_6 &= 0 \\
x_2 + x_6 + x_1 &= 1 \\
x_1 + x_2 + x_5 &= 1 \\
x_3 + x_5 + x_4 &= 0 \qquad \leftarrow \text{Remove} \\
x_6 + x_4 + x_2 &= 1 \qquad \leftarrow \text{Remove}
\end{aligned}
$$

What remains is the 2-core of the system.

The satisfiability threshold is the point where the 2-core has density 1.

The 2-core

Remove every variable that appears in at most one equation, along with the equation it belongs to.

$$
\begin{aligned}
x_5 + x_1 + x_6 &= 0 \\
x_2 + x_6 + x_1 &= 1 \\
x_1 + x_2 + x_5 &= 1 \\
x_3 + x_5 + x_4 &= 0 \qquad \leftarrow \text{Remove} \\
x_6 + x_4 + x_2 &= 1 \qquad \leftarrow \text{Remove}
\end{aligned}
$$

What remains is the 2-core of the system.

Clusters:

- σ is any solution to the 2-core.
- C_σ is all extensions of σ to the rest of the system.

The 2-core

Remove every variable that appears in at most one equation, along with the equation it belongs to.

$$
\begin{aligned}
x_5 + x_1 + x_6 &= 0 \\
x_2 + x_6 + x_1 &= 1 \\
x_1 + x_2 + x_5 &= 1 \\
x_3 + x_5 + x_4 &= 0 \qquad \leftarrow \text{Remove} \\
x_6 + x_4 + x_2 &= 1 \qquad \leftarrow \text{Remove}
\end{aligned}
$$

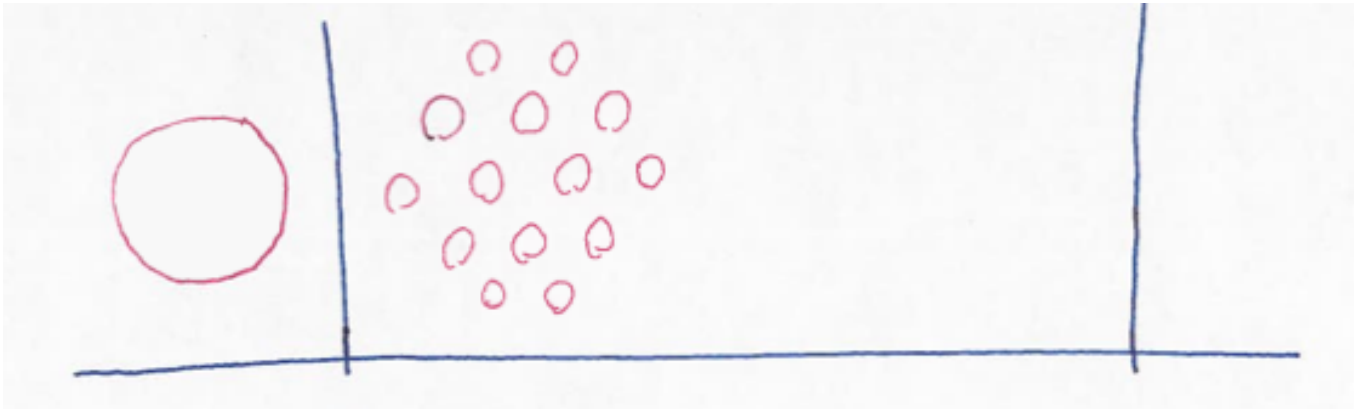
What remains is the 2-core of the system.

Technical correction: We actually need to work with the 2-core minus $O(1)$ variables because of short cycle effects.

Clusters

Clusters:

- σ is any solution to the 2-core.
- C_σ is all extensions of σ to the rest of the system.



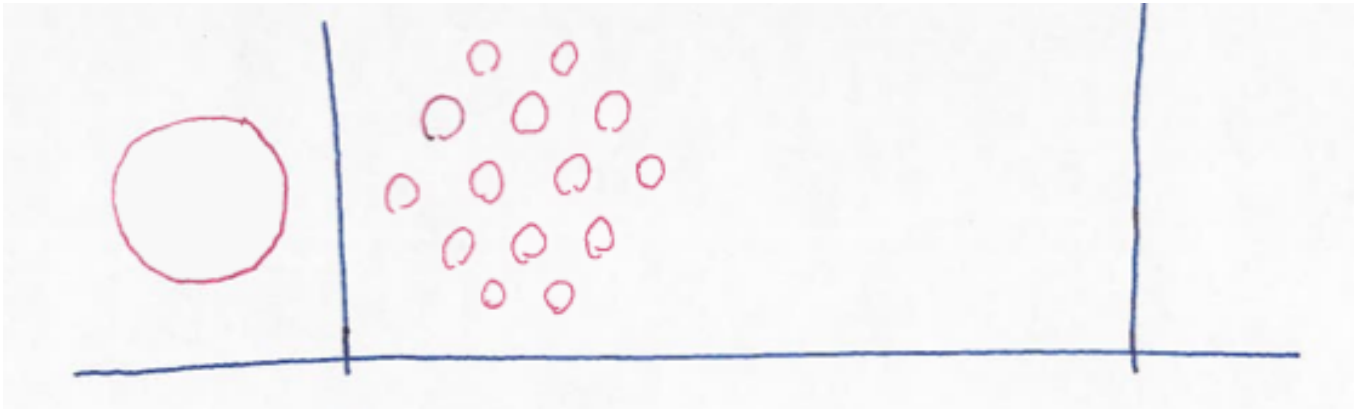
Roughly speaking, clusters are:

- Well-connected. One can move throughout the cluster changing $o(n)$ vertices at a time.
- Well-separated Moving from one cluster to another requires changing $\Theta(n)$ vertices in one step.

Clusters

Clusters:

- σ is any solution to the 2-core.
- C_σ is all extensions of σ to the rest of the system.



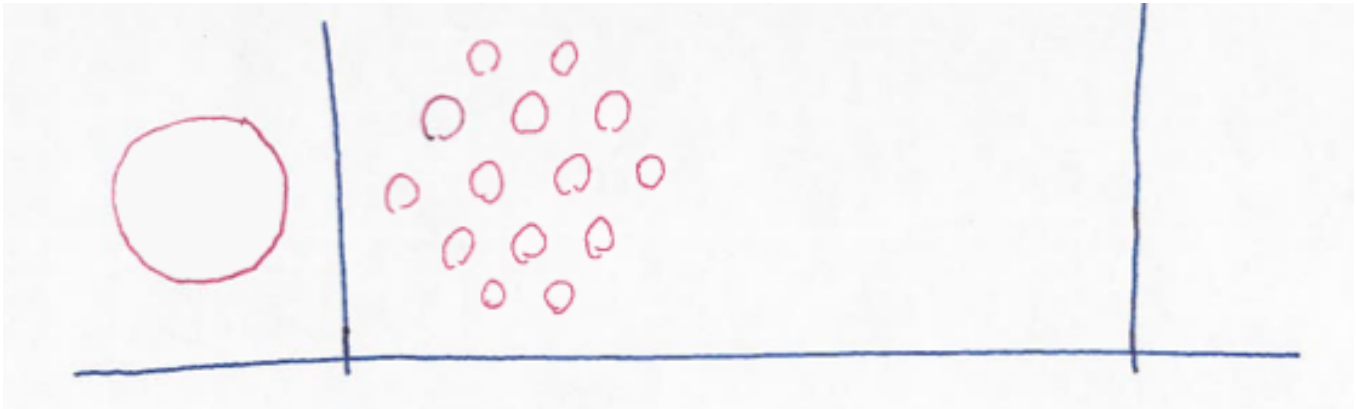
Theorem (IKKM, AM)

- *Any pair of 2-core solutions must differ on at least αn variables.*
- *We can move from any extension of σ to any other extension of σ by changing $O(\log n)$ variables at a time.*

Clusters

Clusters:

- σ is any solution to the 2-core.
- $\mathcal{C}_\sigma$ is all extensions of σ to the rest of the system.

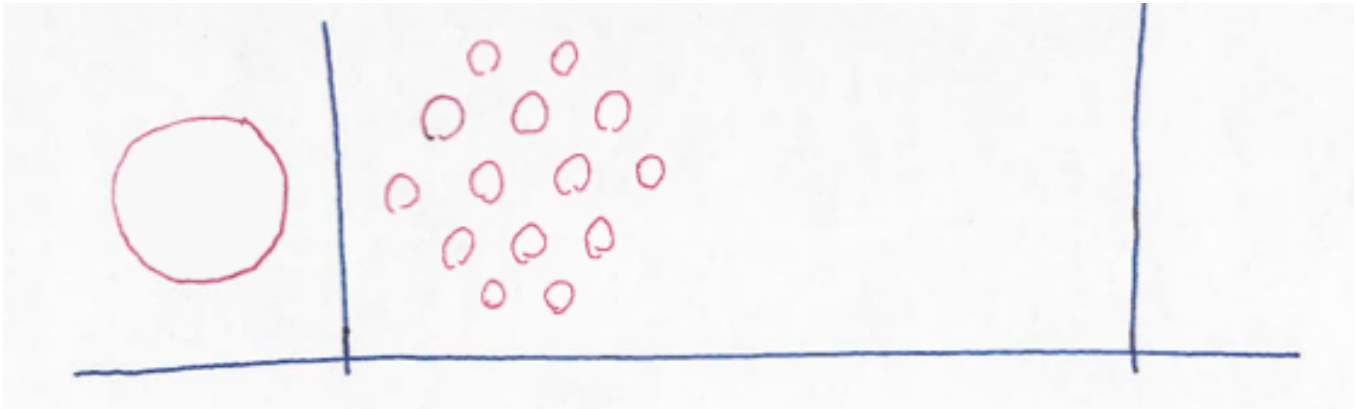


- By symmetry, all clusters are isomorphic.

Clusters

Clusters:

- σ is any solution to the 2-core.
- $\mathcal{C}_\sigma$ is all extensions of σ to the rest of the system.

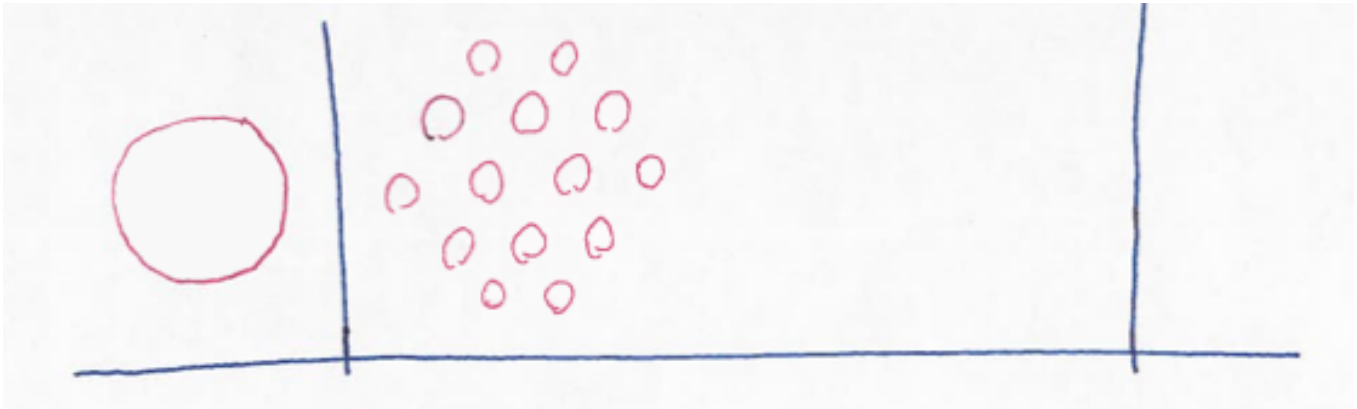


- By symmetry, all clusters are isomorphic.
- The same variables are frozen in every cluster.

Clusters

Clusters:

- σ is any solution to the 2-core.
- $\mathcal{C}_\sigma$ is all extensions of σ to the rest of the system.

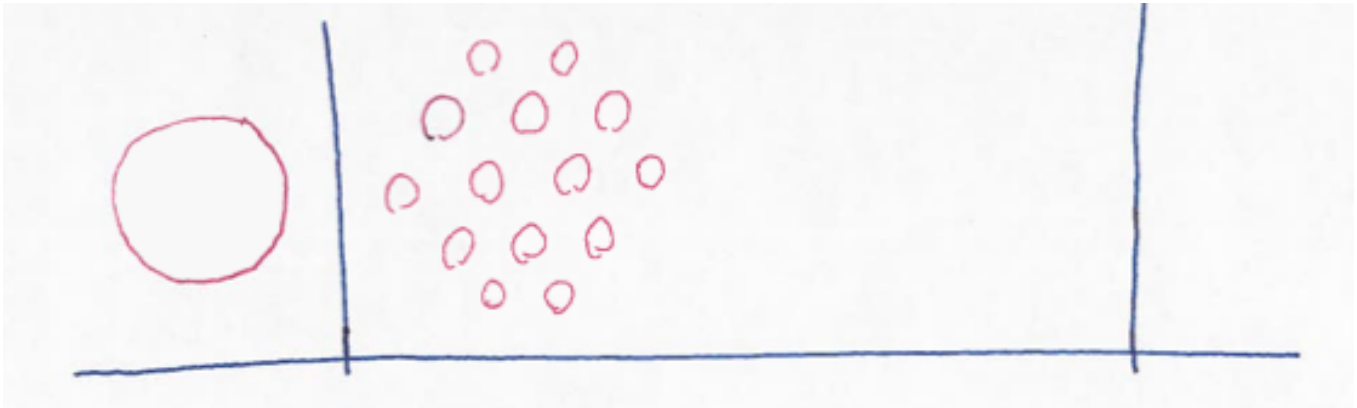


- By symmetry, all clusters are isomorphic.
- The same variables are frozen in every cluster.
- No condensation.

Clusters

Clusters:

- σ is any solution to the 2-core.
- $\mathcal{C}_\sigma$ is all extensions of σ to the rest of the system.



- By symmetry, all clusters are isomorphic.
- The same variables are frozen in every cluster.
- No condensation.
- We only need to analyze the random hypergraph, not actual solutions of the CSP.

Proven For Other CSPs

Well-Separated Clusters: Asymptotic in k threshold for

- k-SAT, k-colouring, hypergraph 2-colouring Achlioptas and Coja-Oghlin 2008
- independent set Coja-Oghlin and Efthymiou 2010
- several others Montanari, Restrepo, Tetali 2009

Proven For Other CSPs

Well-Separated Clusters: Asymptotic in k threshold for

- k-SAT, k-colouring, hypergraph 2-colouring Achlioptas and Coja-Oghlin 2008
- independent set Coja-Oghlin and Efthymiou 2010
- several others Montanari, Restrepo, Tetali 2009

Freezing:

- occurs in k-SAT Achlioptas and Ricci-Tersinghi 2006
- Asymptotic in k threshold for k-SAT, k-colouring, hypergraph 2-colouring Achlioptas and Coja-Oghlin 2008
- exact threshold for k-colouring M 2011
- exact threshold for hypergraph 2-colouring and others M and Restrepo 2013

Gaussian Elimination

Stripping Digraph: When we remove x, we direct an edge from x to each of the other $k-1$ variables in its equation.

Gaussian Elimination

Stripping Digraph: When we remove x, we direct an edge from x to each of the other $k-1$ variables in its equation.

Gaussian Elimination: Eliminate variables as they are removed.

$$x_i = \sum_{x_j \in \chi_i} x_j$$

Gaussian Elimination

Stripping Digraph: When we remove x, we direct an edge from x to each of the other $k-1$ variables in its equation.

Gaussian Elimination: Eliminate variables as they are removed.

$$x_i = \sum_{x_j \in \chi_i} x_j$$

Easy: x_i can reach every $x_j \in \chi_i$.

Gaussian Elimination

Stripping Digraph: When we remove x, we direct an edge from x to each of the other $k-1$ variables in its equation.

Gaussian Elimination: Eliminate variables as they are removed.

$$x_i = \sum_{x_j \in \chi_i} x_j$$

Easy: x_i can reach every $x_j \in \chi_i$.

Lemma

Each non 2-core x_j is reachable from $O(\log n)$ other variables.

So we can move between solutions in C_σ by changing base variables one at a time. Each change affects $O(\log n)$ variables.

Depth in r-cores

r-core of a hypergraph: Repeatedly remove vertices of degree $< r$.

Note that the order in which vertices are removed does not affect the core obtained.

This is the largest subgraph with minimum degree at least r.

Depth in r-cores

r-core of a hypergraph: Repeatedly remove vertices of degree $< r$.

Note that the order in which vertices are removed does not affect the core obtained.

This is the largest subgraph with minimum degree at least r.

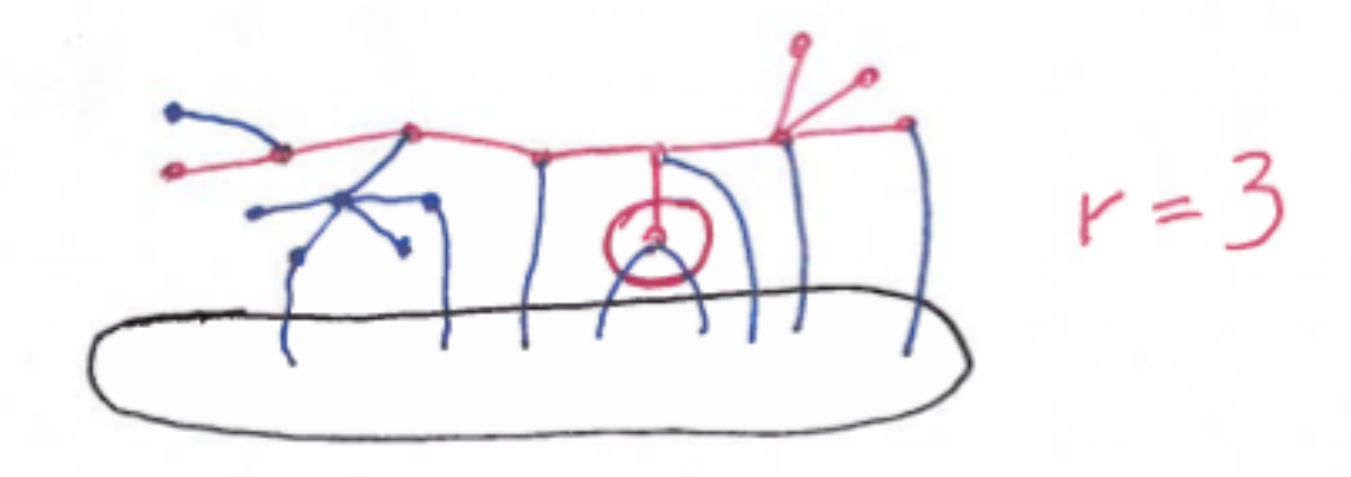
Analyzed for random graphs in Pittel, Spencer, Wormald(1996) and many other papers.

Applications include colouring, hashing, coding, orientability,...

Depth in r-cores

r-core of a hypergraph: Repeatedly remove vertices of degree $< r$.

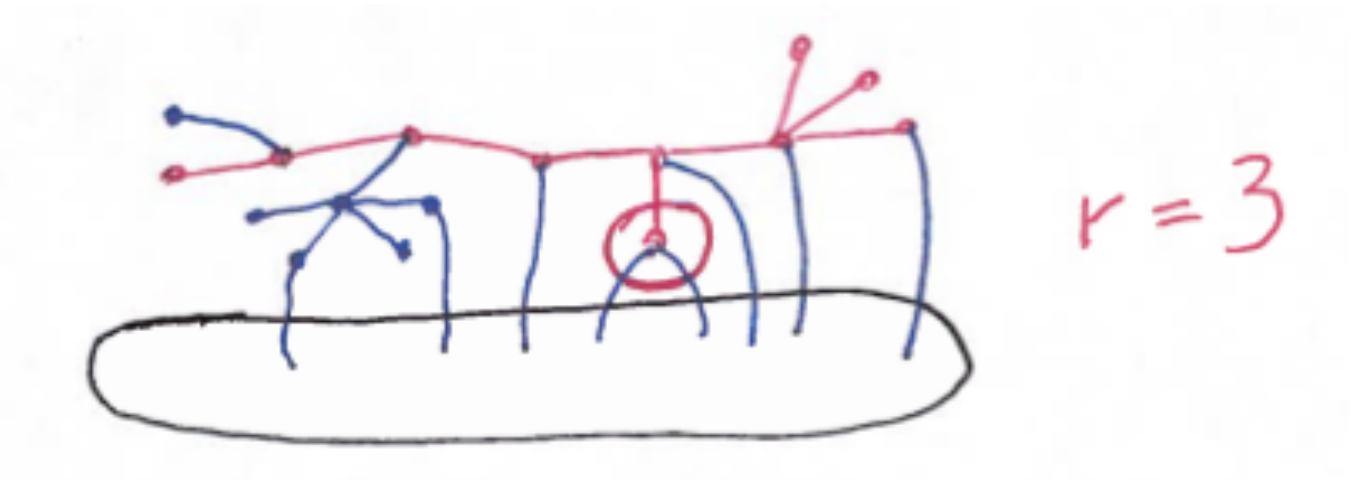
The depth of a non r-core vertex v is the shortest sequence of deletions that leads to removing v.



Depth in r-cores

r-core of a hypergraph: Repeatedly remove vertices of degree $< r$.

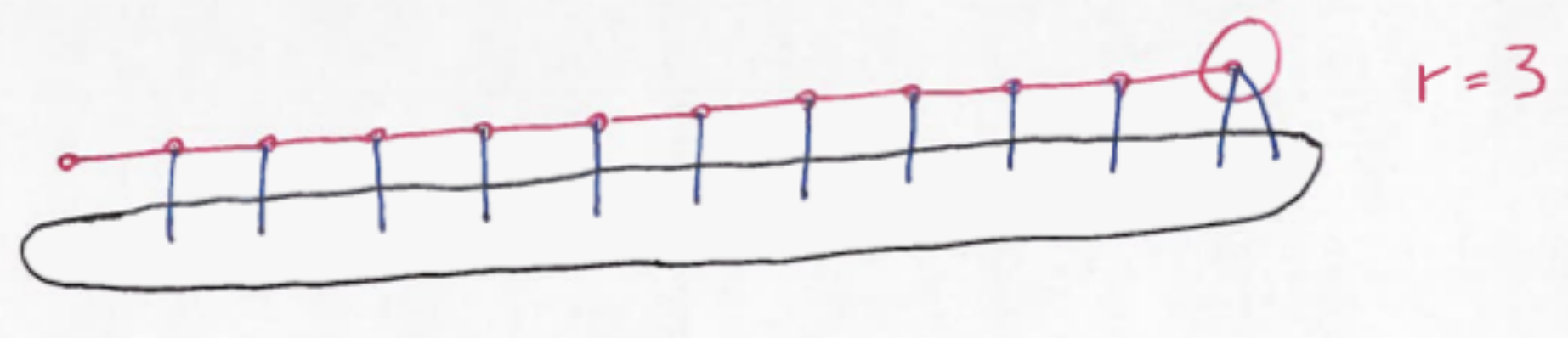
The depth of a non r-core vertex v is the shortest sequence of deletions that leads to removing v.



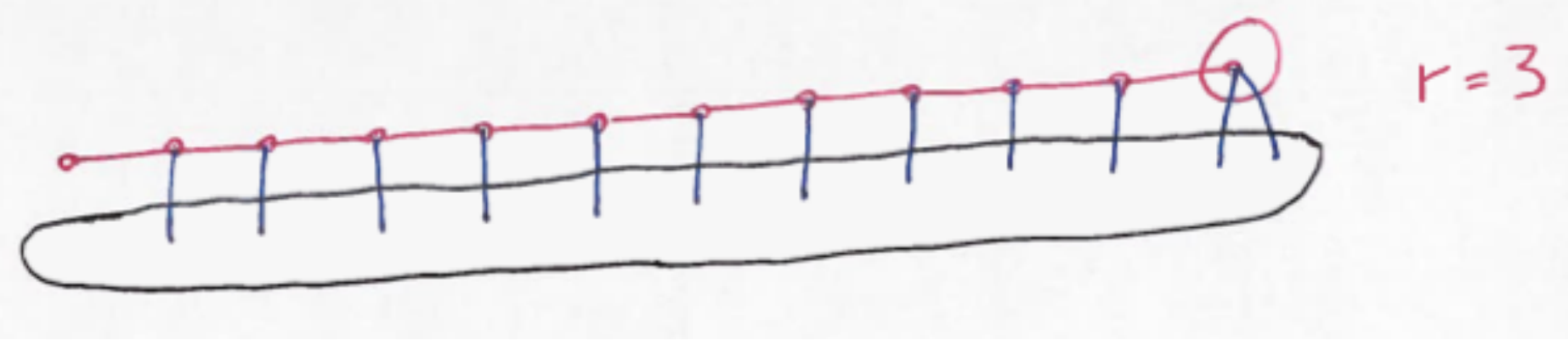
Theorem (Achlioptas and M)

For any $c \neq c^*$ (the r-core threshold), the maximum depth in $H^k(n, M = cn)$ is $O(\log n)$.

Key Step



Key Step



When we remove a vertex, the expected number of degree r neighbours is at most $1 - \epsilon$.

Inside the Clustering Threshold

c^* is the clustering threshold.

Inside the Clustering Threshold

c^* is the clustering threshold.

Theorem (IKKM, AM 2011)

- *Any pair of 2-core solutions must differ on at least αn variables.*
- *We can move from any extension of σ to any other extension of σ by changing $O(\log n)$ variables at a time.*

Inside the Clustering Threshold

c^* is the clustering threshold.

Theorem (Gao and M)

For sufficiently small $\delta > 0$, and $c = c^* + n^{-\delta}$:

- *Any pair of 2-core solutions must differ on at least $n^{1-\beta}$ variables.*
- *We can move from any extension of σ to any other extension of σ by changing n^{β} variables at a time.*

($\beta \to 0$ with δ.)

Inside the Clustering Threshold

c^* is the clustering threshold.

Theorem (Gao and M)

For sufficiently small $\delta > 0$, and $c = c^* + n^{-\delta}$:

- *Any pair of 2-core solutions must differ on at least $n^{1-\beta}$ variables.*
- *We can move from any extension of σ to any other extension of σ by changing n^{β} variables at a time.*

($\beta \to 0$ with δ.)

Theorem (Achlioptas and M)

For any $c \neq c^*$, the maximum depth in $H^k(n, M = cn)$ is $O(\log n)$.

Inside the Clustering Threshold

c^* is the clustering threshold.

Theorem (Gao and M)

For sufficiently small $\delta > 0$, and $c = c^* + n^{-\delta}$:

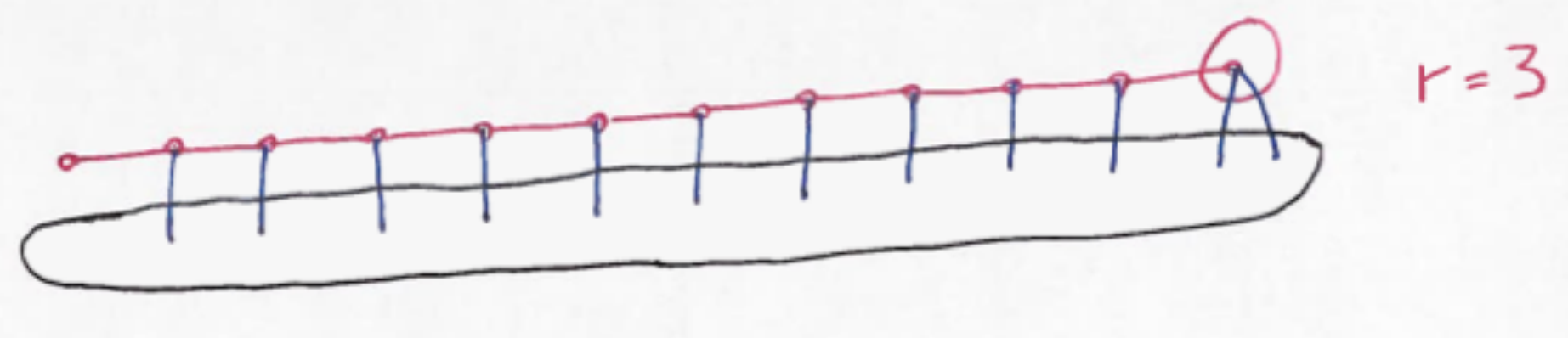
- *Any pair of 2-core solutions must differ on at least $n^{1-\beta}$ variables.*
- *We can move from any extension of σ to any other extension of σ by changing n^{β} variables at a time.*

($\beta \to 0$ with δ.)

Theorem (Gao and M)

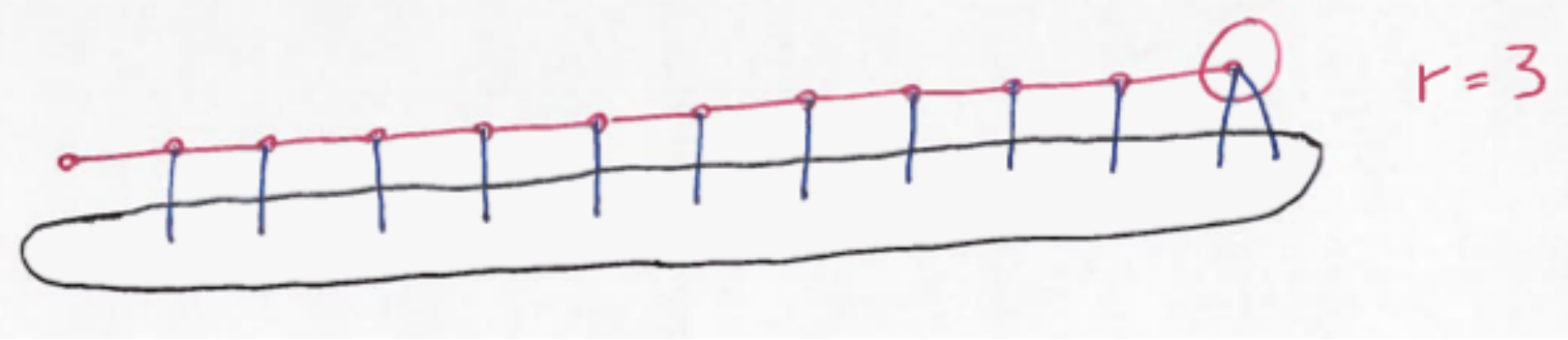
For sufficiently small $\delta > 0$, and $c = c^* + n^{-\delta}$: the maximum depth in $H^k(n, M = cn)$ is at most n^{β}.

Challenge



When we remove a vertex, the expected number of degree r neighbours is at most $1 - \epsilon$.

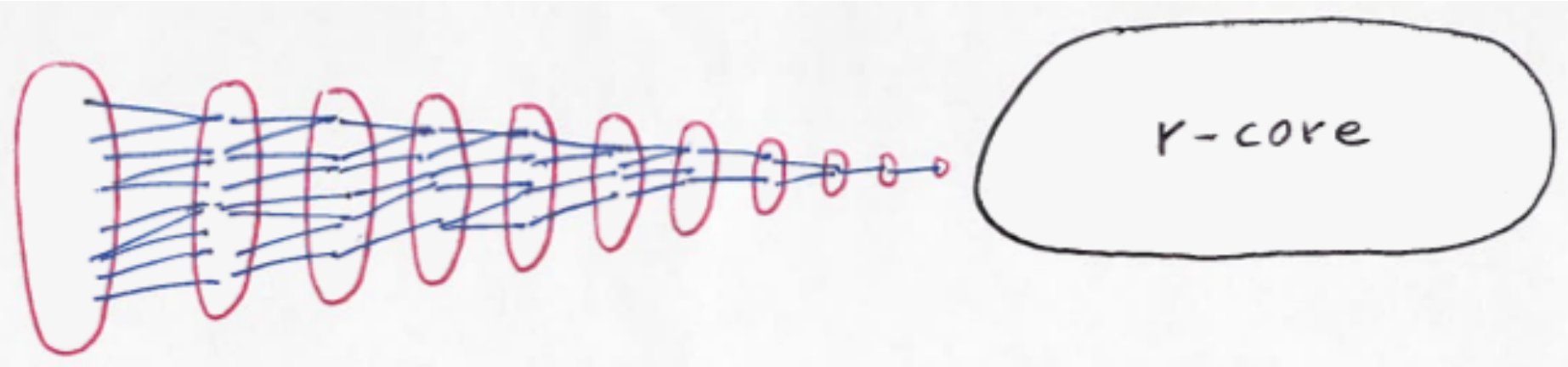
Challenge



When we remove a vertex, the expected number of degree r neighbours approaches 1, as we approach the r-core.

Parallel Stripping

Parallel Stripping Process: At each iteration, simultaneously remove every vertex of degree less than r.

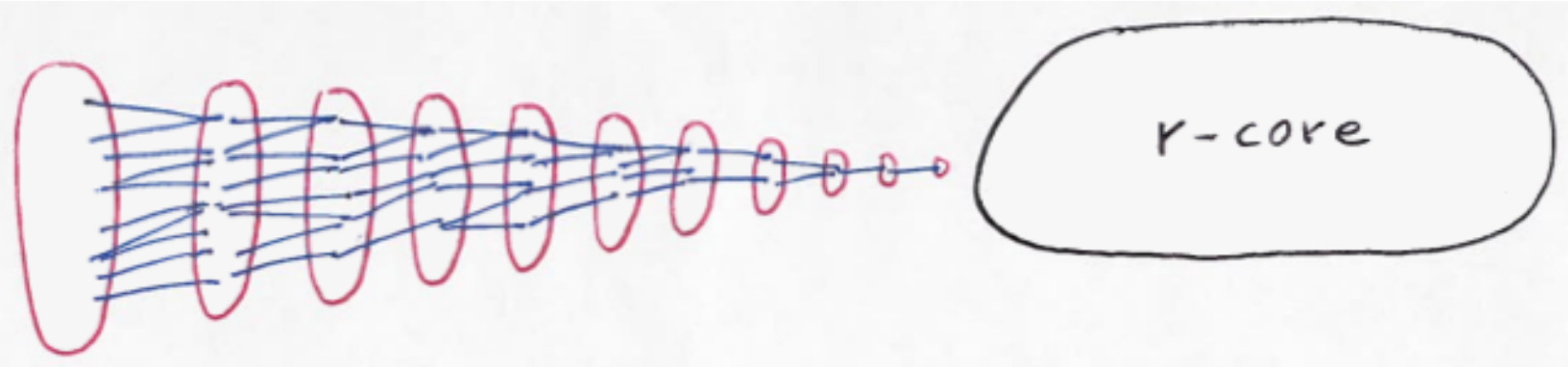


We begin by determining the number of iterations:

$$c = c^* + n^{-\delta} \qquad \rightarrow \qquad \approx n^{\delta/2} \text{ iterations}$$

Parallel Stripping

Parallel Stripping Process: At each iteration, simultaneously remove every vertex of degree less than r.



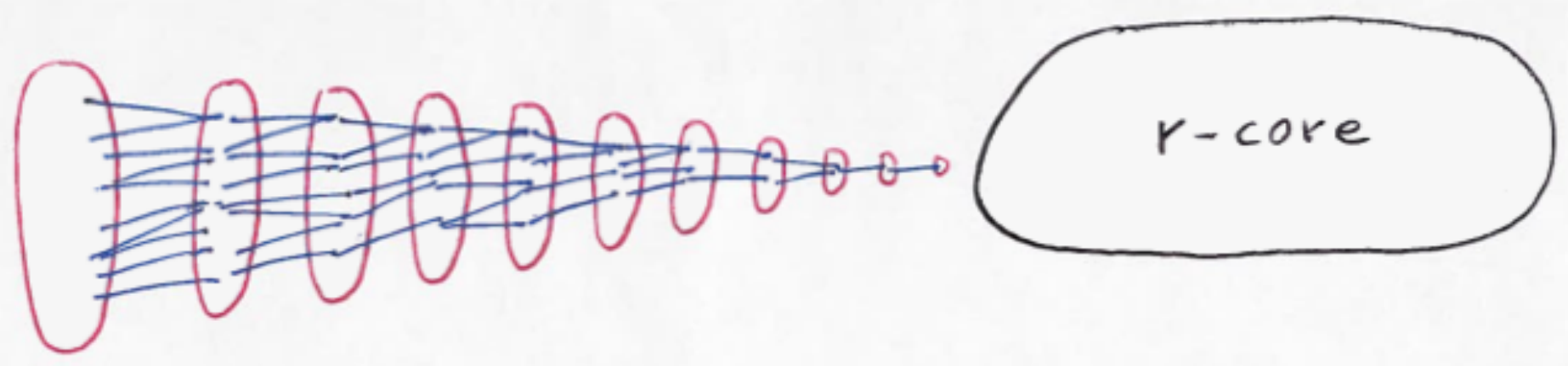
We begin by determining the number of iterations:

$$c = c^* + n^{-\delta} \qquad \rightarrow \qquad \approx n^{\delta/2} \text{ iterations}$$

This lower bounds the maximum depth.

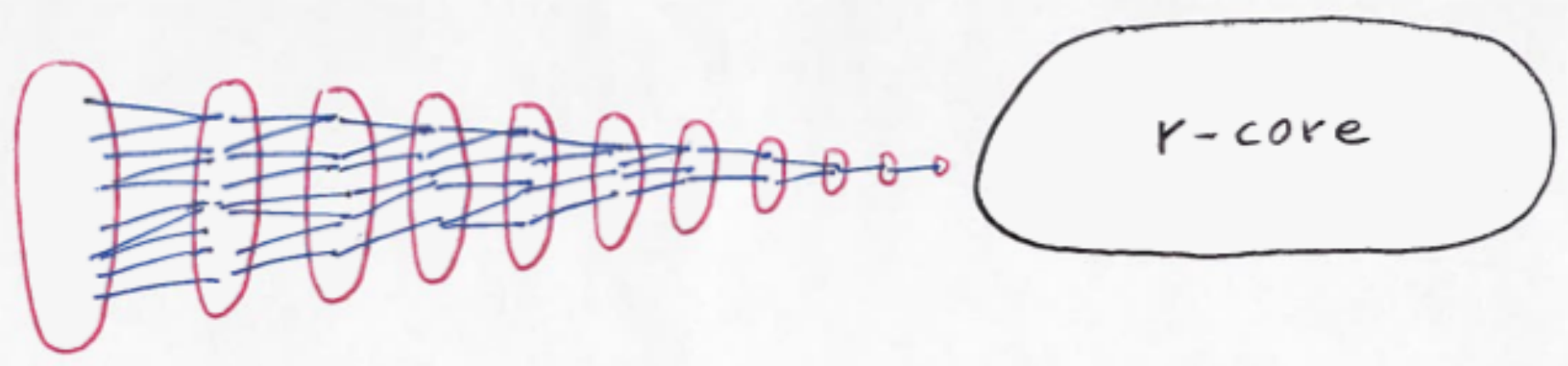
Edge Switching

We carry out the parallel stripping process.



Edge Switching

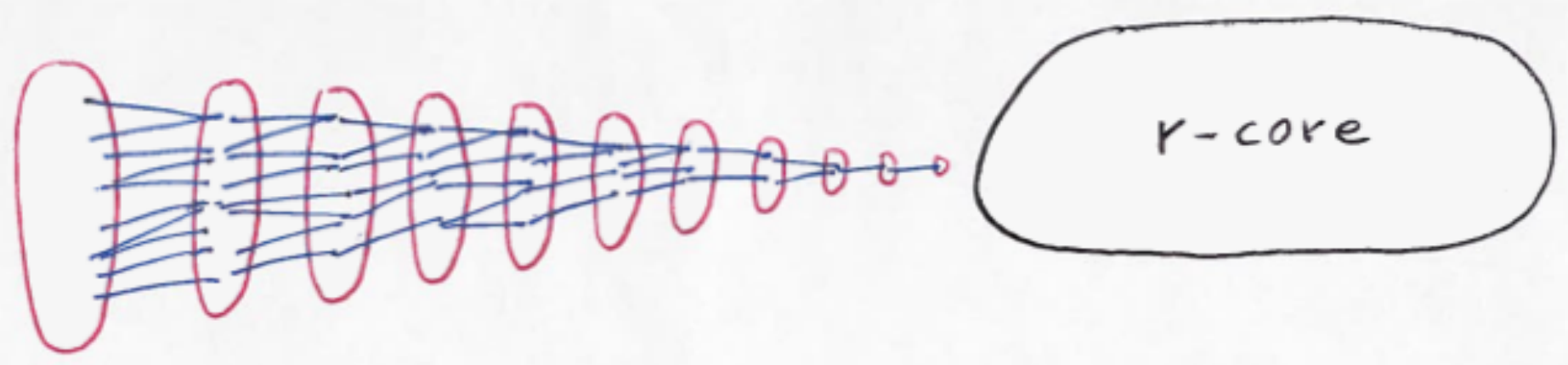
We carry out the parallel stripping process.



Then we randomly switch the blue edges.

Edge Switching

We carry out the parallel stripping process.

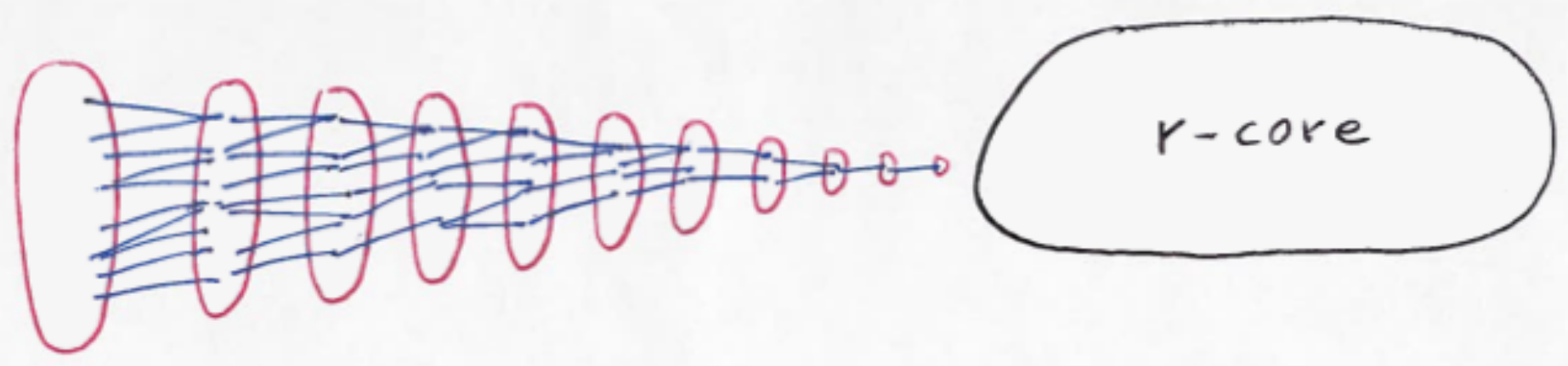


Then we randomly switch the blue edges.

Equivalently - we first expose the vertices deleted in each iteration, but not the edges between them.

Edge Switching

We carry out the parallel stripping process.



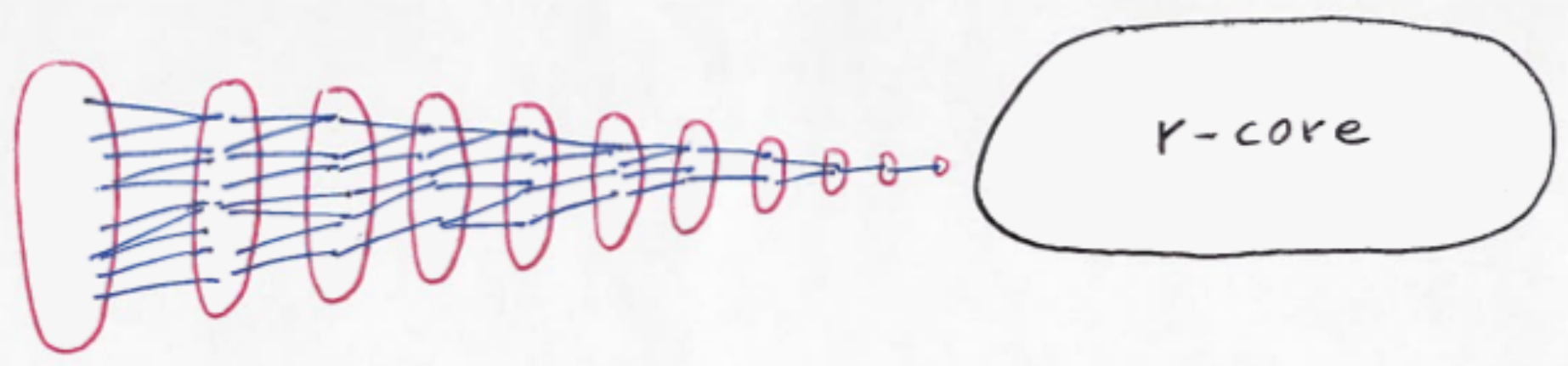
Then we randomly switch the blue edges.

Subject to conditions such as:

- Every vertex as at least one neighbour in the preceding level.

Edge Switching

We carry out the parallel stripping process.



Then we randomly switch the blue edges.

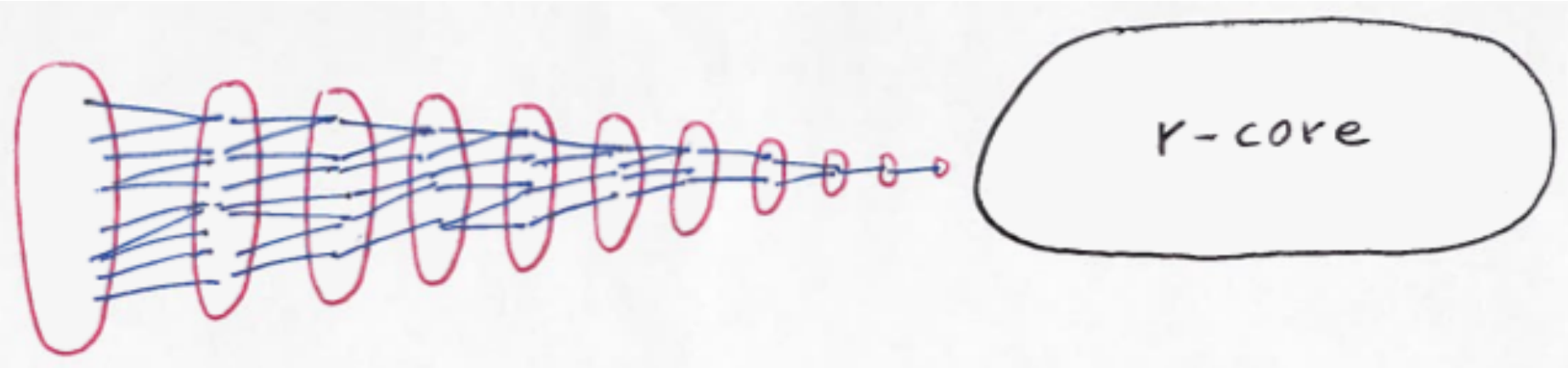
Subject to conditions such as:

- Every vertex as at least one neighbour in the preceding level.

This provides enough randomness for us to analyze the stripping sequence leading to a particular vertex being removed.

Big Steps Required Inside the Clusters

Within each C_σ, changing some variables requires changing $n^{O(\delta)}$ others.



Further Challenges

- Push further into the clustering threshold.
- Gain a better rigorous understanding of clusters for other CSP's.