The small subgraph conditioning method and hypergraphs

Catherine Greenhill

School of Mathematics and Statistics
UNSW Sydney

The small subgraph conditioning method:

The small subgraph conditioning method:

An analysis of variance technique introduced by Robinson & Wormald (1992).

The small subgraph conditioning method:

An analysis of variance technique introduced by Robinson & Wormald (1992).

Technique for analysing a random variable $Y = Y_n$, particularly to show that $\Pr(Y > 0) \to 1$, where the second moment method does not apply.

The small subgraph conditioning method:

An analysis of variance technique introduced by Robinson & Wormald (1992).

Technique for analysing a random variable $Y = Y_n$, particularly to show that $\Pr(Y > 0) \to 1$, where the second moment method does not apply.

Also establishes the asymptotic distribution of Y and a property called contiguity of two probability spaces.

The small subgraph conditioning method:

An analysis of variance technique introduced by Robinson & Wormald (1992).

Technique for analysing a random variable $Y = Y_n$, particularly to show that $\Pr(Y > 0) \to 1$, where the second moment method does not apply.

Also establishes the asymptotic distribution of Y and a property called contiguity of two probability spaces.

See Wormald's 1999 regular graphs survey
+ Janson's 1995 paper with "contiguity" in the title.

Suppose we have a sequence of probability spaces $\mathcal{G}_n$ indexed by n, and a random variable $Y = Y_n$ defined on $\mathcal{G}_n$.

Suppose we have a sequence of probability spaces $\mathcal{G}_n$ indexed by n, and a random variable $Y = Y_n$ defined on $\mathcal{G}_n$.

Want to show $Y > 0$ asymptotically almost surely (a.a.s.); that is, $\Pr(Y > 0) \to 1$ as $n \to \infty$.

Suppose we have a sequence of probability spaces $\mathcal{G}_n$ indexed by n, and a random variable $Y = Y_n$ defined on $\mathcal{G}_n$.

Want to show $Y > 0$ asymptotically almost surely (a.a.s.); that is, $\Pr(Y > 0) \to 1$ as $n \to \infty$.

If $\mathbb{E}(Y^2) = (1 + o(1))\, \mathbb{E}(Y)^2$ then, by Chebyshev's inequality,

$$\Pr(Y = 0) \le \Pr(|Y - \mathbb{E}\,Y| \ge \mathbb{E}\,Y) \le \frac{\mathrm{Var}(Y)}{(\mathbb{E}\,Y)^2}$$

Suppose we have a sequence of probability spaces $\mathcal{G}_n$ indexed by n, and a random variable $Y = Y_n$ defined on $\mathcal{G}_n$.

Want to show $Y > 0$ asymptotically almost surely (a.a.s.); that is, $\Pr(Y > 0) \to 1$ as $n \to \infty$.

If $\mathbb{E}(Y^2) = (1 + o(1))\, \mathbb{E}(Y)^2$ then, by Chebyshev's inequality,

$$\begin{aligned} \Pr(Y = 0) \le \Pr(|Y - \mathbb{E}\, Y| \ge \mathbb{E}\, Y) &\le \frac{\mathrm{Var}(Y)}{(\mathbb{E}\, Y)^2} \\ &= \frac{\mathbb{E}(Y^2)}{(\mathbb{E}\, Y)^2} - 1 \end{aligned}$$

Suppose we have a sequence of probability spaces $\mathcal{G}_n$ indexed by n, and a random variable $Y = Y_n$ defined on $\mathcal{G}_n$.

Want to show $Y > 0$ asymptotically almost surely (a.a.s.); that is, $\Pr(Y > 0) \to 1$ as $n \to \infty$.

If $\mathbb{E}(Y^2) = (1 + o(1))\, \mathbb{E}(Y)^2$ then, by Chebyshev's inequality,

$$\begin{aligned}\Pr(Y = 0) \le \Pr(|Y - \mathbb{E}\, Y| \ge \mathbb{E}\, Y) &\le \frac{\mathrm{Var}(Y)}{(\mathbb{E}\, Y)^2} \\ &= \frac{\mathbb{E}(Y^2)}{(\mathbb{E}\, Y)^2} - 1 = o(1).\end{aligned}$$

Suppose we have a sequence of probability spaces $\mathcal{G}_n$ indexed by n, and a random variable $Y = Y_n$ defined on $\mathcal{G}_n$.

Want to show $Y > 0$ asymptotically almost surely (a.a.s.); that is, $\Pr(Y > 0) \to 1$ as $n \to \infty$.

If $\mathbb{E}(Y^2) = (1 + o(1))\, \mathbb{E}(Y)^2$ then, by Chebyshev's inequality,

$$\Pr(Y = 0) \leq \Pr(|Y - \mathbb{E}\,Y| \geq \mathbb{E}\,Y) \leq \frac{\mathrm{Var}(Y)}{(\mathbb{E}\,Y)^2} = \frac{\mathbb{E}(Y^2)}{(\mathbb{E}\,Y)^2} - 1 = o(1).$$

Second moment method works.

What if $\frac{\mathbb{E}(Y^2)}{(\mathbb{E}\,Y)^2} \to C$ for some constant $C > 1$?

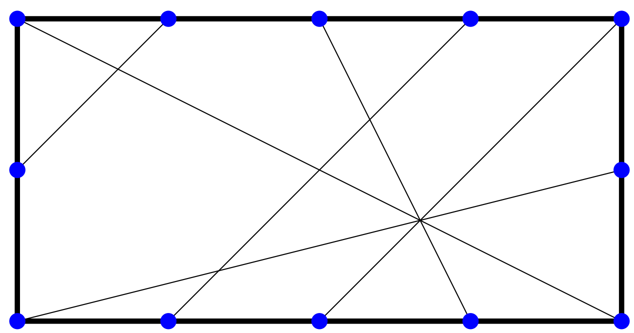
What if $\frac{\mathbb{E}(Y^2)}{(\mathbb{E}Y)^2} \to C$ for some constant $C > 1$?

The second moment method is not strong enough.

What if $\frac{\mathbb{E}(Y^2)}{(\mathbb{E}\,Y)^2} \to C$ for some constant $C > 1$?

The second moment method is not strong enough.

Robinson & Wormald faced exactly this problem when studying Hamilton cycles in random 3-regular graphs.



Write $[n] = \{1, 2, \ldots, n\}$ and let $\mathcal{G}_{n,d}$ denote a uniformly random d-regular graph on $[n]$, where d is fixed.

Write $[n] = \{1, 2, \ldots, n\}$ and let $\mathcal{G}_{n,d}$ denote a uniformly random d-regular graph on $[n]$, where d is fixed.

In 1984, Robinson & Wormald proved $\frac{\mathbb{E}(Y^2)}{(\mathbb{E}\, Y)^2} \to 3/e$, where Y is the number of Hamilton cycles in $\mathcal{G}_{n,3}$.

Write $[n] = \{1, 2, \ldots, n\}$ and let $\mathcal{G}_{n,d}$ denote a uniformly random d-regular graph on $[n]$, where d is fixed.

In 1984, Robinson & Wormald proved $\frac{\mathbb{E}(Y^2)}{(\mathbb{E}\,Y)^2} \to 3/e$, where Y is the number of Hamilton cycles in $\mathcal{G}_{n,3}$.

This implied that $\Pr(Y > 0) \geq 2 - 3/e \approx 0.896$.

Write $[n] = \{1, 2, \ldots, n\}$ and let $\mathcal{G}_{n,d}$ denote a uniformly random d-regular graph on $[n]$, where d is fixed.

In 1984, Robinson & Wormald proved $\frac{\mathbb{E}(Y^2)}{(\mathbb{E}\, Y)^2} \to 3/e$, where Y is the number of Hamilton cycles in $\mathcal{G}_{n,3}$.

This implied that $\Pr(Y > 0) \geq 2 - 3/e \approx 0.896$.

In the same paper, Robinson & Wormald improved this to

$$\Pr(Y > 0) \geq 2 - 3/e^{13/12} \approx 0.985$$

by studying triangle-free 3-regular graphs.

Write $[n] = \{1, 2, \ldots, n\}$ and let $\mathcal{G}_{n,d}$ denote a uniformly random d-regular graph on $[n]$, where d is fixed.

In 1984, Robinson & Wormald proved $\frac{\mathbb{E}(Y^2)}{(\mathbb{E}\,Y)^2} \to 3/e$, where Y is the number of Hamilton cycles in $\mathcal{G}_{n,3}$.

This implied that $\Pr(Y > 0) \geq 2 - 3/e \approx 0.896$.

In the same paper, Robinson & Wormald improved this to

$$\Pr(Y > 0) \geq 2 - 3/e^{13/12} \approx 0.985$$

by studying triangle-free 3-regular graphs.

$\Rightarrow$ Small cycles can have a **big** effect!

Robinson & Wormald (1992): Proved that $\Pr(Y > 0) \to 1$, so almost all cubic graphs are Hamiltonian. They wrote that this result "has been suspected for some time".

This paper introduced the small subgraph conditioning method.

Robinson & Wormald (1992): Proved that $\Pr(Y > 0) \to 1$, so almost all cubic graphs are Hamiltonian. They wrote that this result "has been suspected for some time".

This paper introduced the small subgraph conditioning method.

Janson (1995) observed that R & W's proof technique also

- gives the asymptotic distribution of Y,

Robinson & Wormald (1992): Proved that $\Pr(Y > 0) \to 1$, so almost all cubic graphs are Hamiltonian. They wrote that this result "has been suspected for some time".

This paper introduced the small subgraph conditioning method.

Janson (1995) observed that R & W's proof technique also

- gives the asymptotic distribution of Y, and
- establishes a property called "contiguity" between $\mathcal{G}_{n,3}$ and a probability space, denoted $\mathcal{G}_{n,3}^{(Y)}$, where each 3-regular graph G on $[n]$ has probability proportional to $Y(G)$.

Le Cam (1960):

Suppose $(\mathcal{A}_n)$ and $(\mathcal{B}_n)$ are two sequences of probability spaces on the same sequence of underlying sets (Ω_n).

Le Cam (1960):

Suppose $(\mathcal{A}_n)$ and $(\mathcal{B}_n)$ are two sequences of probability spaces on the same sequence of underlying sets (Ω_n).

Say $(\mathcal{A}_n)$ and $(\mathcal{B}_n)$ are (mutually) contiguous if

$$\mathsf{Pr}_{\mathcal{A}_n}(E_n) \to 1 \text{ if and only if } \mathsf{Pr}_{\mathcal{B}_n}(E_n) \to 1$$

for all $E_n \subseteq \Omega_n$.

Le Cam (1960):

Suppose $(\mathcal{A}_n)$ and $(\mathcal{B}_n)$ are two sequences of probability spaces on the same sequence of underlying sets (Ω_n).

Say $(\mathcal{A}_n)$ and $(\mathcal{B}_n)$ are (mutually) contiguous if

$$\mathrm{Pr}_{\mathcal{A}_n}(E_n) \to 1 \text{ if and only if } \mathrm{Pr}_{\mathcal{B}_n}(E_n) \to 1$$

for all $E_n \subseteq \Omega_n$.

Write $\mathcal{A}_n \approx \mathcal{B}_n$ when $(\mathcal{A}_n)$ and $(\mathcal{B}_n)$ are contiguous.

Janson: contiguity is "qualitative asymptotic equivalence".

Recall, $\mathcal{G}_{n,3}^{(Y)}$ gives each 3-regular graph G on $[n]$ probability proportional to $Y(G)$.

Recall, $\mathcal{G}_{n,3}^{(Y)}$ gives each 3-regular graph G on $[n]$ probability proportional to $Y(G)$.

Robinson & Wormald (1992), restated: $\mathcal{G}_{n,3}^{(Y)} \approx \mathcal{G}_{n,3}$.

Recall, $\mathcal{G}_{n,3}^{(Y)}$ gives each 3-regular graph G on $[n]$ probability proportional to $Y(G)$.

Robinson & Wormald (1992), restated: $\mathcal{G}_{n,3}^{(Y)} \approx \mathcal{G}_{n,3}$.

Observe, a graph in $\mathcal{G}_{n,3}^{(Y)}$ is Hamiltonian with probability 1.

Recall, $\mathcal{G}_{n,3}^{(Y)}$ gives each 3-regular graph G on $[n]$ probability proportional to $Y(G)$.

Robinson & Wormald (1992), restated: $\mathcal{G}_{n,3}^{(Y)} \approx \mathcal{G}_{n,3}$.

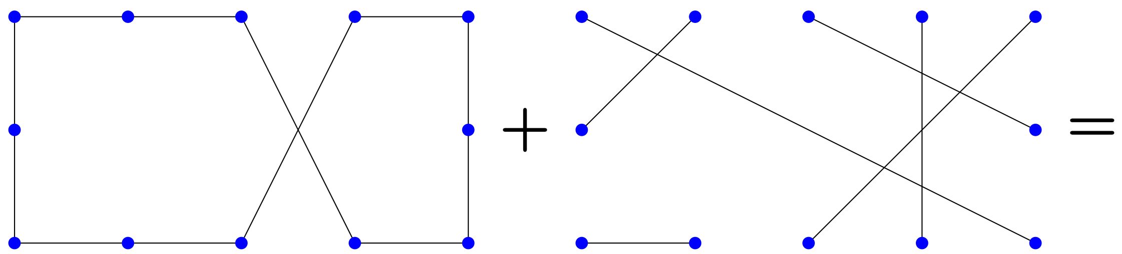
Observe, a graph in $\mathcal{G}_{n,3}^{(Y)}$ is Hamiltonian with probability 1.

Then contiguity immediately implies that

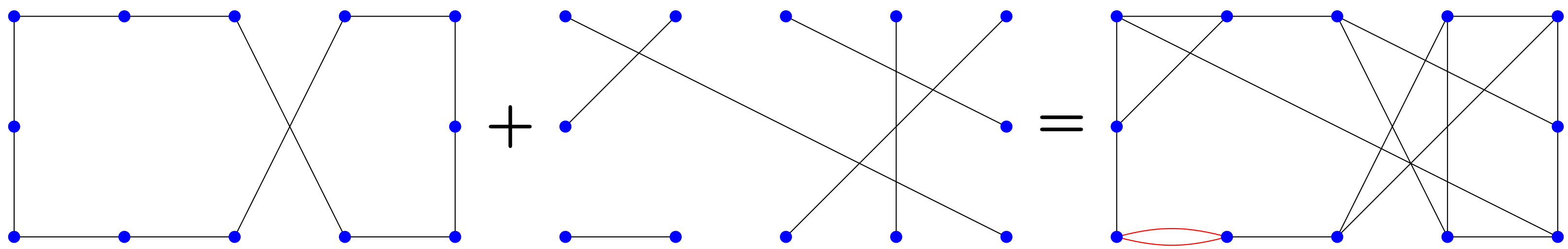
$$\Pr(\mathcal{G}_{n,3} \text{ is Hamiltonian}) \to 1.$$

Can also rephrase R & W result to say: $\mathcal{G}_{n,3}$ is contiguous with the superposition of a uniformly random Hamilton cycle and a uniformly random perfect matching, both on $[n]$.

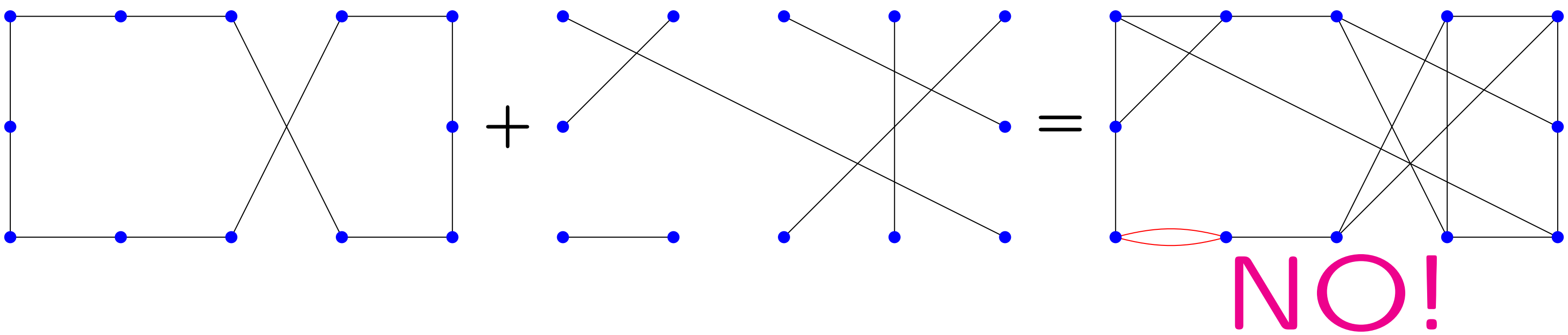
Can also rephrase R & W result to say: $\mathcal{G}_{n,3}$ is contiguous with the superposition of a uniformly random Hamilton cycle and a uniformly random perfect matching, both on $[n]$.



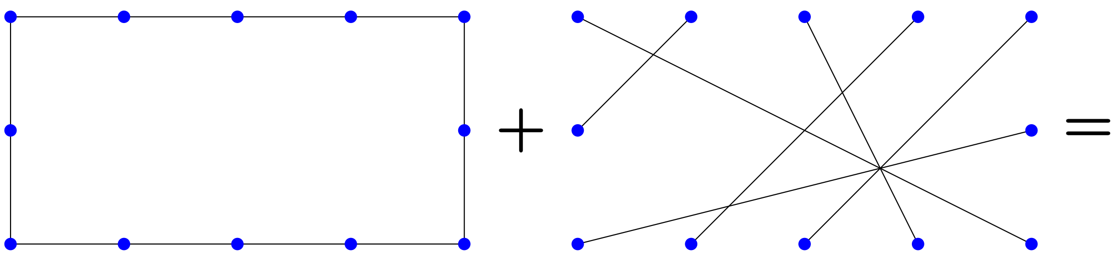
Can also rephrase R & W result to say: $\mathcal{G}_{n,3}$ is contiguous with the superposition of a uniformly random Hamilton cycle and a uniformly random perfect matching, both on $[n]$.



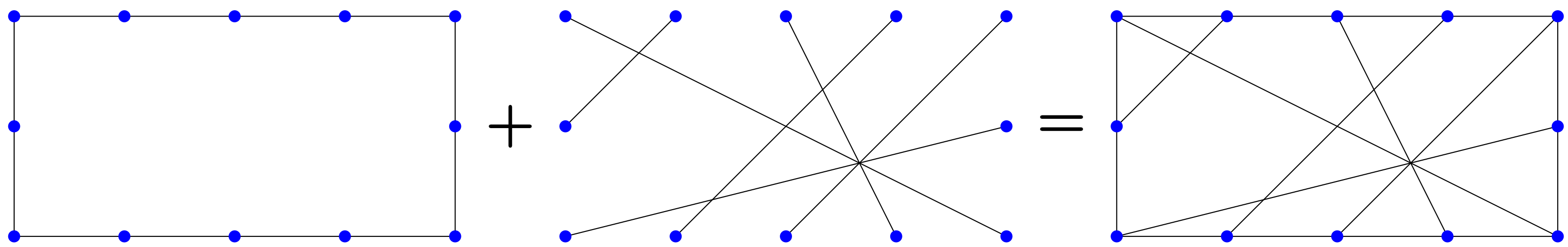
Can also rephrase R & W result to say: $\mathcal{G}_{n,3}$ is contiguous with the superposition of a uniformly random Hamilton cycle and a uniformly random perfect matching, both on $[n]$.



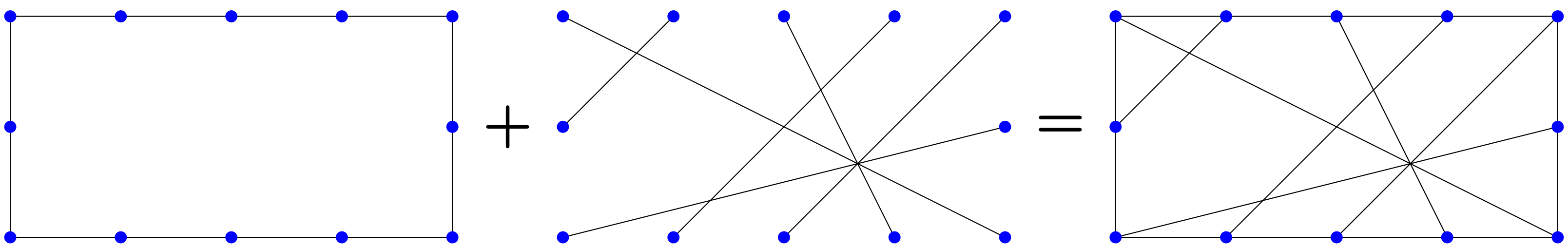
Can also rephrase R & W result to say: $\mathcal{G}_{n,3}$ is contiguous with the superposition of a uniformly random Hamilton cycle and a uniformly random perfect matching, both on $[n]$.



Can also rephrase R & W result to say: $\mathcal{G}_{n,3}$ is contiguous with the superposition of a uniformly random Hamilton cycle and a uniformly random perfect matching, both on $[n]$.

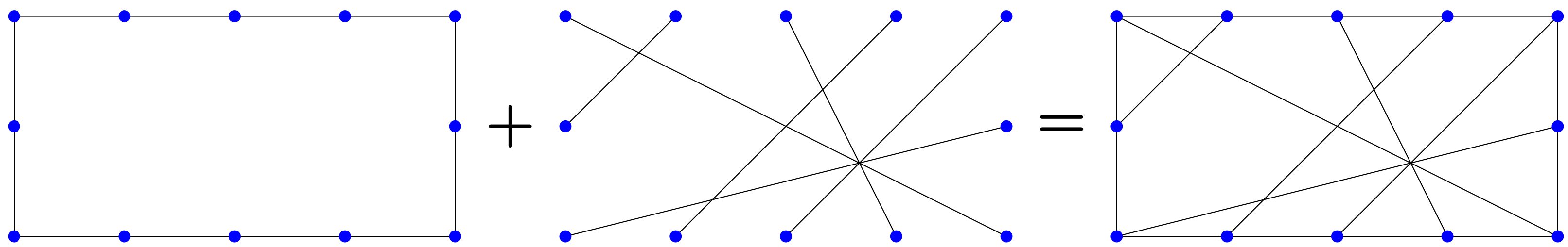


Can also rephrase R & W result to say: $\mathcal{G}_{n,3}$ is contiguous with the superposition of a uniformly random Hamilton cycle and a uniformly random perfect matching, both on $[n]$.



Leads to "contiguity arithmetic". See Wormald's 1999 survey.

Can also rephrase R & W result to say: $\mathcal{G}_{n,3}$ is contiguous with the superposition of a uniformly random Hamilton cycle and a uniformly random perfect matching, both on $[n]$.



Leads to "contiguity arithmetic". See Wormald's 1999 survey. (Warning: $1 + 1 \neq 2$.)

Small subgraph conditioning method (SSCM)

Robinson & Wormald (1992,1994), Janson (1995)

Small subgraph conditioning method (SSCM)

Robinson & Wormald (1992,1994), Janson (1995)

Let $\lambda_i > 0$ and $\delta_i \geq -1$ be constants, for $i \geq 1$.

Small subgraph conditioning method (SSCM)
Robinson & Wormald (1992,1994), Janson (1995)

Let $\lambda_i > 0$ and $\delta_i \geq -1$ be constants, for $i \geq 1$.

Suppose that for all n you have random variables X_{in} and Y_n, defined on same probability space $\mathcal{G}_n$, where the X_{in} are nonnegative integer-valued and $\mathbb{E}\, Y_n \neq 0$.

Small subgraph conditioning method (SSCM)
Robinson & Wormald (1992,1994), Janson (1995)

Let $\lambda_i > 0$ and $\delta_i \geq -1$ be constants, for $i \geq 1$.

Suppose that for all n you have random variables X_{in} and Y_n, defined on same probability space $\mathcal{G}_n$, where the X_{in} are nonnegative integer-valued and $\mathbb{E}\, Y_n \neq 0$.

Further suppose that:

(A1) $X_{in} \xrightarrow{d} Z_i$ as $n \to \infty$, jointly for all $i \geq 1$, where $Z_i \sim \mathrm{Po}(\lambda_i)$ are independent Poisson.

(A2) For any sequence $x_1, \ldots, x_m$ of nonnegative integers,

$$\frac{\mathbb{E}(Y_n \mid X_{1n} = x_1, \ldots, X_{mn} = x_m)}{\mathbb{E}\, Y_n} \to \prod_{i=1}^{m} (1 + \delta_i)^{x_i}\, e^{-\lambda_i \delta_i}$$

as $n \to \infty$.

(A2) For any sequence $x_1, \ldots, x_m$ of nonnegative integers,

$$\frac{\mathbb{E}(Y_n \mid X_{1n} = x_1, \ldots, X_{mn} = x_m)}{\mathbb{E}\, Y_n} \to \prod_{i=1}^{m} (1 + \delta_i)^{x_i}\, e^{-\lambda_i \delta_i}$$

as $n \to \infty$.

(A3) $\sum_{i=1}^{\infty} \lambda_i \delta_i^2 < \infty$.

(A2) For any sequence $x_1, \ldots, x_m$ of nonnegative integers,

$$\frac{\mathbb{E}(Y_n \mid X_{1n} = x_1, \ldots, X_{mn} = x_m)}{\mathbb{E}\, Y_n} \to \prod_{i=1}^{m} (1 + \delta_i)^{x_i}\, e^{-\lambda_i \delta_i}$$

as $n \to \infty$.

(A3) $\sum_{i=1}^{\infty} \lambda_i \delta_i^2 < \infty$.

(A4) $\frac{\mathbb{E}(Y_n^2)}{(\mathbb{E}\, Y_n)^2} \to \exp\left(\sum_{i=1}^{\infty} \lambda_i \delta_i^2\right)$ as $n \to \infty$.

(A2) For any sequence $x_1, \ldots, x_m$ of nonnegative integers,

$$\frac{\mathbb{E}(Y_n \mid X_{1n} = x_1, \ldots, X_{mn} = x_m)}{\mathbb{E}\, Y_n} \to \prod_{i=1}^{m} (1+\delta_i)^{x_i}\, e^{-\lambda_i \delta_i}$$

as $n \to \infty$.

(A3) $\sum_{i=1}^{\infty} \lambda_i \delta_i^2 < \infty$.

(A4) $\frac{\mathbb{E}(Y_n^2)}{(\mathbb{E}\, Y_n)^2} \to \exp\left(\sum_{i=1}^{\infty} \lambda_i \delta_i^2\right)$ as $n \to \infty$.

Then (distribution version, Janson 1995):

$$\frac{Y_n}{\mathbb{E}\, Y_n} \xrightarrow{d} W = \prod_{i=1}^{\infty} (1+\delta_i)^{Z_i}\, e^{-\lambda_i \delta_i} \quad \text{as } n \to \infty;$$

(A2) For any sequence $x_1, \dots, x_m$ of nonnegative integers,

$$\frac{\mathbb{E}(Y_n \mid X_{1n} = x_1, \dots, X_{mn} = x_m)}{\mathbb{E}\, Y_n} \to \prod_{i=1}^{m} (1+\delta_i)^{x_i} \, e^{-\lambda_i \delta_i}$$

as $n \to \infty$.

(A3) $\sum_{i=1}^{\infty} \lambda_i \delta_i^2 < \infty$.

(A4) $\frac{\mathbb{E}(Y_n^2)}{(\mathbb{E}\, Y_n)^2} \to \exp\left(\sum_{i=1}^{\infty} \lambda_i \delta_i^2\right)$ as $n \to \infty$.

Then (distribution version, Janson 1995):

$$\frac{Y_n}{\mathbb{E}\, Y_n} \xrightarrow{d} W = \prod_{i=1}^{\infty} (1+\delta_i)^{Z_i} \, e^{-\lambda_i \delta_i} \quad \text{as } n \to \infty;$$

moreover, this and the convergence in (A1) hold jointly.

(A2) For any sequence $x_1, \ldots, x_m$ of nonnegative integers,

$$\frac{\mathbb{E}(Y_n \mid X_{1n} = x_1, \ldots, X_{mn} = x_m)}{\mathbb{E}\, Y_n} \to \prod_{i=1}^{m} (1 + \delta_i)^{x_i} \, e^{-\lambda_i \delta_i}$$

as $n \to \infty$.

(A3) $\sum_{i=1}^{\infty} \lambda_i \delta_i^2 < \infty$.

(A4) $\frac{\mathbb{E}(Y_n^2)}{(\mathbb{E}\, Y_n)^2} \to \exp\left(\sum_{i=1}^{\infty} \lambda_i \delta_i^2\right)$ as $n \to \infty$.

Then (distribution version, Janson 1995):

$$\frac{Y_n}{\mathbb{E}\, Y_n} \xrightarrow{d} W = \prod_{i=1}^{\infty} (1 + \delta_i)^{Z_i} \, e^{-\lambda_i \delta_i} \quad \text{as } n \to \infty;$$

moreover, this and the convergence in (A1) hold jointly.
Also, $W > 0$ almost surely if and only if $\delta_i > -1$ for all i.

(A2) For any sequence $x_1, \ldots, x_m$ of nonnegative integers,

$$\frac{\mathbb{E}(Y_n \mid X_{1n} = x_1, \ldots, X_{mn} = x_m)}{\mathbb{E}\, Y_n} \to \prod_{i=1}^{m} (1 + \delta_i)^{x_i}\, e^{-\lambda_i \delta_i}$$

as $n \to \infty$.

(A3) $\sum_{i=1}^{\infty} \lambda_i \delta_i^2 < \infty$.

(A4) $\frac{\mathbb{E}(Y_n^2)}{(\mathbb{E}\, Y_n)^2} \to \exp\left(\sum_{i=1}^{\infty} \lambda_i \delta_i^2\right)$ as $n \to \infty$.

(A2) For any sequence $x_1, \ldots, x_m$ of nonnegative integers,

$$\frac{\mathbb{E}(Y_n \mid X_{1n} = x_1, \ldots, X_{mn} = x_m)}{\mathbb{E}\, Y_n} \to \prod_{i=1}^{m} (1 + \delta_i)^{x_i} e^{-\lambda_i \delta_i}$$

as $n \to \infty$.

(A3) $\sum_{i=1}^{\infty} \lambda_i \delta_i^2 < \infty$.

(A4) $\frac{\mathbb{E}(Y_n^2)}{(\mathbb{E}\, Y_n)^2} \to \exp\left(\sum_{i=1}^{\infty} \lambda_i \delta_i^2\right)$ as $n \to \infty$.

Then (contiguity version, Wormald 1999):

$$\Pr(Y_n > 0) = \exp\left(- \sum_{\delta_i = -1} \lambda_i \right) + o(1)$$

and $\mathcal{G}_n^{(Y_n)} \approx \mathcal{G}_n$ if $\delta_i > -1$ for all i.

Many structural results about regular graphs, regular bipartite graphs, proved using SSCM by various authors:

Delcourt, Frieze, Greenhill, Janson, Jerrum, Kim, Kwan, Molloy, Postle, Prałat, Robalewska, Robinson, Ruciński, Shi, Wind, Wormald.

(Apologies to any I've missed!)

Many structural results about regular graphs, regular bipartite graphs, proved using SSCM by various authors:

Delcourt, Frieze, Greenhill, Janson, Jerrum, Kim, Kwan, Molloy, Postle, Prałat, Robalewska, Robinson, Ruciński, Shi, Wind, Wormald.

(Apologies to any I've missed!)

A couple of examples:

Kim & Wormald (2001)

A.a.s. $\mathcal{G}_{n,4}$ is the union of two Hamilton cycles.

Many structural results about regular graphs, regular bipartite graphs, proved using SSCM by various authors:

Delcourt, Frieze, Greenhill, Janson, Jerrum, Kim, Kwan, Molloy, Postle, Prałat, Robalewska, Robinson, Ruciński, Shi, Wind, Wormald.

(Apologies to any I've missed!)

A couple of examples:

Kim & Wormald (2001)

A.a.s. $\mathcal{G}_{n,4}$ is the union of two Hamilton cycles.

Prałat & Wormald (2019)

Almost all 5-regular graphs have a 3-flow.

More recent applications use SSCM to study thresholds in random constraint satisfaction problems, e.g.

More recent applications use SSCM to study thresholds in random constraint satisfaction problems, e.g.

Bapst, Coja-Oghlan, Efthymiou (2017),
random colourings in $G(n, m)$

Coja-Oghlan & Wormald (2018), random k-SAT formulae

Coja-Oghlan, Kapetanopoulos, Müller (2020+),
random constraint satisfaction problems

More recent applications use SSCM to study thresholds in random constraint satisfaction problems, e.g.

Bapst, Coja-Oghlan, Efthymiou (2017),
random colourings in $G(n, m)$

Coja-Oghlan & Wormald (2018), random k-SAT formulae

Coja-Oghlan, Kapetanopoulos, Müller (2020+),
random constraint satisfaction problems

Idea: the planted model is easier to study, so prove that this model is contiguous with respect to the standard model.

More recent applications use SSCM to study thresholds in random constraint satisfaction problems, e.g.

Bapst, Coja-Oghlan, Efthymiou (2017),
random colourings in $G(n, m)$

Coja-Oghlan & Wormald (2018), random k-SAT formulae

Coja-Oghlan, Kapetanopoulos, Müller (2020+),
random constraint satisfaction problems

Idea: the planted model is easier to study, so prove that this model is contiguous with respect to the standard model.

Also: Bank, Moore, Neeman, Netrapalli (2016),
community detection in sparse networks.

Rest of talk: structural results for regular graphs, or regular uniform hypergraphs. Firstly, regular graphs.

Rest of talk: structural results for regular graphs, or regular uniform hypergraphs. Firstly, regular graphs.

Usually, calculations are performed in the configuration model where X_{in} is the number of cycles of length i.

Rest of talk: structural results for regular graphs, or regular uniform hypergraphs. Firstly, regular graphs.

Usually, calculations are performed in the configuration model where X_{in} is the number of cycles of length i.

(These are the “small subgraphs” in the name of the method.)

Rest of talk: structural results for regular graphs, or regular uniform hypergraphs. Firstly, regular graphs.

Usually, calculations are performed in the configuration model where X_{in} is the number of cycles of length i.

(These are the "small subgraphs" in the name of the method.)

Bollobás (1980): the X_{in} are asymptotically independent Poisson with mean $(d-1)^i/(2i)$.

Rest of talk: structural results for regular graphs, or regular uniform hypergraphs. Firstly, regular graphs.

Usually, calculations are performed in the configuration model where X_{in} is the number of cycles of length i.

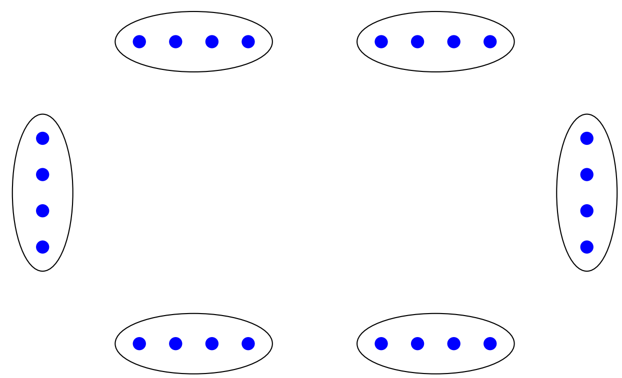
(These are the "small subgraphs" in the name of the method.)

Bollobás (1980): the X_{in} are asymptotically independent Poisson with mean $(d-1)^i/(2i)$.

The SSCM works when the variance of Y is well-controlled by the short cycle counts.

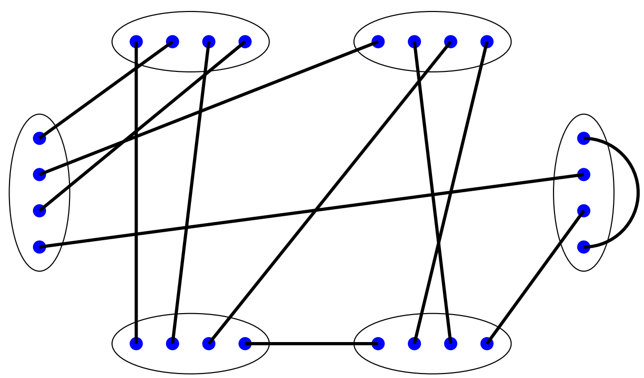
Configuration model (Bollobás, 1980)

Start with n cells, each containing d points. Take a uniformly random perfect matching of dn points into $dn/2$ pairs.



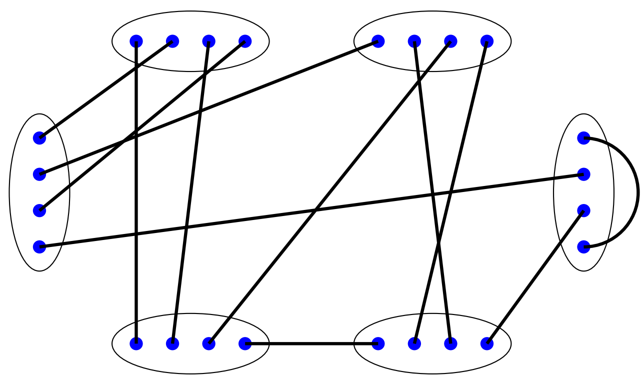
Configuration model (Bollobás, 1980)

Start with n cells, each containing d points. Take a uniformly random perfect matching of dn points into $dn/2$ pairs.



Configuration model (Bollobás, 1980)

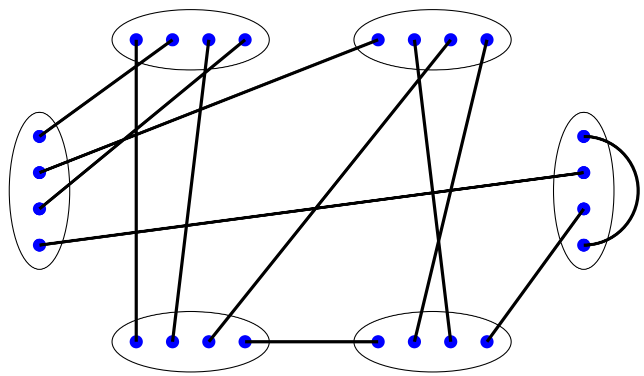
Start with n cells, each containing d points. Take a uniformly random perfect matching of dn points into $dn/2$ pairs.



Shrink each cell to a vertex to get a d-regular multigraph. Every simple graph is equally likely.

Configuration model (Bollobás, 1980)

Start with n cells, each containing d points. Take a uniformly random perfect matching of dn points into $dn/2$ pairs.



Shrink each cell to a vertex to get a d-regular multigraph. Every simple graph is equally likely.

Bender & Canfield (1978): Pr(simple) $\sim e^{-(d^2-1)/4}$.

Suppose that Y is a random variable of interest in a random configuration, and $Y_{\mathcal{G}}$ is the corresponding variable in $\mathcal{G}_{n,d}$.

Suppose that Y is a random variable of interest in a random configuration, and $Y_{\mathcal{G}}$ is the corresponding variable in $\mathcal{G}_{n,d}$.

Assume we have proved (A1)–(A4) hold for Y.

If $\Pr(Y = 0) = o(1)$ then

$$\Pr(Y_{\mathcal{G}} = 0) \leq \frac{\Pr(Y = 0)}{\Pr(\text{simple})} = o(1).$$

Suppose that Y is a random variable of interest in a random configuration, and $Y_{\mathcal{G}}$ is the corresponding variable in $\mathcal{G}_{n,d}$.

Assume we have proved (A1)–(A4) hold for Y.

If $\Pr(Y = 0) = o(1)$ then

$$\Pr(Y_{\mathcal{G}} = 0) \leq \frac{\Pr(Y = 0)}{\Pr(\text{simple})} = o(1).$$

Also, if the uniform and Y-weighted configuration models are contiguous then $\mathcal{G}_{n,d}^{(Y)} \approx \mathcal{G}_{n,d}$.

Suppose that Y is a random variable of interest in a random configuration, and $Y_{\mathcal{G}}$ is the corresponding variable in $\mathcal{G}_{n,d}$.

Suppose that Y is a random variable of interest in a random configuration, and $Y_{\mathcal{G}}$ is the corresponding variable in $\mathcal{G}_{n,d}$.

Assume we have proved (A1)–(A4) hold for Y, and hence

$$\frac{Y}{\mathbb{E}\,Y} \xrightarrow{d} W = \prod_{i=1}^{\infty} (1+\delta_i)^{Z_i}\, e^{-\lambda_i \delta_i}$$

for some constants λ_i, δ_i, and where the random variables $Z_i \sim \mathsf{Po}(\lambda_i)$ are independent.

Suppose that Y is a random variable of interest in a random configuration, and $Y_{\mathcal{G}}$ is the corresponding variable in $\mathcal{G}_{n,d}$.

Assume we have proved (A1)–(A4) hold for Y, and hence

$$\frac{Y}{\mathbb{E}\,Y} \xrightarrow{d} W = \prod_{i=1}^{\infty} (1+\delta_i)^{Z_i}\, e^{-\lambda_i \delta_i}$$

for some constants λ_i, δ_i, and where the random variables $Z_i \sim \mathrm{Po}(\lambda_i)$ are independent.

Now (A2) implies that

$$\frac{\mathbb{E}\,Y_{\mathcal{G}}}{\mathbb{E}\,Y} = \frac{\mathbb{E}(Y \mid X_{1n} = X_{2n} = 0)}{\mathbb{E}\,Y} \to e^{-\lambda_1\delta_1 - \lambda_2\delta_2}.$$

(A configuration gives a simple graph iff $X_{1n} = X_{2n} = 0$.)

By the joint convergence of $(\frac{Y}{\mathbb{E}(Y)}, X_{1n}, X_{2n})$ to (W, Z_1, Z_2), we conclude that

$$\begin{aligned}\mathcal{L}\left(\frac{Y_{\mathcal{G}}}{\mathbb{E}\,Y}\right) &= \mathcal{L}\left(\frac{Y}{\mathbb{E}\,Y}\middle| X_{1n} = X_{2n} = 0\right) \\ &\xrightarrow{d} \mathcal{L}(W \mid Z_1 = Z_2 = 0) \\ &= \mathcal{L}\left(e^{-\lambda_1\delta_1 - \lambda_2\delta_2} \prod_{i=3}^{\infty} (1+\delta_i)^{Z_i}\, e^{-\lambda_i\delta_i}\right).\end{aligned}$$

By the joint convergence of $(\frac{Y}{\mathbb{E}(Y)}, X_{1n}, X_{2n})$ to (W, Z_1, Z_2), we conclude that

$$\begin{aligned}
\mathcal{L}\left(\frac{Y_{\mathcal{G}}}{\mathbb{E}\,Y}\right) &= \mathcal{L}\left(\frac{Y}{\mathbb{E}\,Y}\middle| X_{1n} = X_{2n} = 0\right) \\
&\xrightarrow{d} \mathcal{L}(W \mid Z_1 = Z_2 = 0) \\
&= \mathcal{L}\left(e^{-\lambda_1\delta_1 - \lambda_2\delta_2} \prod_{i=3}^{\infty} (1+\delta_i)^{Z_i}\, e^{-\lambda_i\delta_i}\right).
\end{aligned}$$

Hence

$$\frac{Y_{\mathcal{G}}}{\mathbb{E}\,Y_{\mathcal{G}}} \sim e^{\lambda_1\delta_1 + \lambda_2\delta_2} \frac{Y_{\mathcal{G}}}{\mathbb{E}\,Y} \xrightarrow{d} \prod_{i=3}^{\infty} (1+\delta_i)^{Z_i}\, e^{-\lambda_i\delta_i}.$$

By the joint convergence of $(\frac{Y}{\mathbb{E}(Y)}, X_{1n}, X_{2n})$ to (W, Z_1, Z_2), we conclude that

$$\mathcal{L}\left(\frac{Y_{\mathcal{G}}}{\mathbb{E}\,Y}\right) = \mathcal{L}\left(\frac{Y}{\mathbb{E}\,Y}\middle| X_{1n} = X_{2n} = 0\right)$$
$$\xrightarrow{d} \mathcal{L}(W \mid Z_1 = Z_2 = 0)$$
$$= \mathcal{L}\left(e^{-\lambda_1\delta_1 - \lambda_2\delta_2} \prod_{i=3}^{\infty} (1+\delta_i)^{Z_i}\, e^{-\lambda_i\delta_i}\right).$$

Hence

$$\frac{Y_{\mathcal{G}}}{\mathbb{E}\,Y_{\mathcal{G}}} \sim e^{\lambda_1\delta_1 + \lambda_2\delta_2} \frac{Y_{\mathcal{G}}}{\mathbb{E}\,Y} \xrightarrow{d} \prod_{i=3}^{\infty} (1+\delta_i)^{Z_i}\, e^{-\lambda_i\delta_i}.$$

TL;DR Delete $i = 1, 2$ factors to get result for regular graphs!

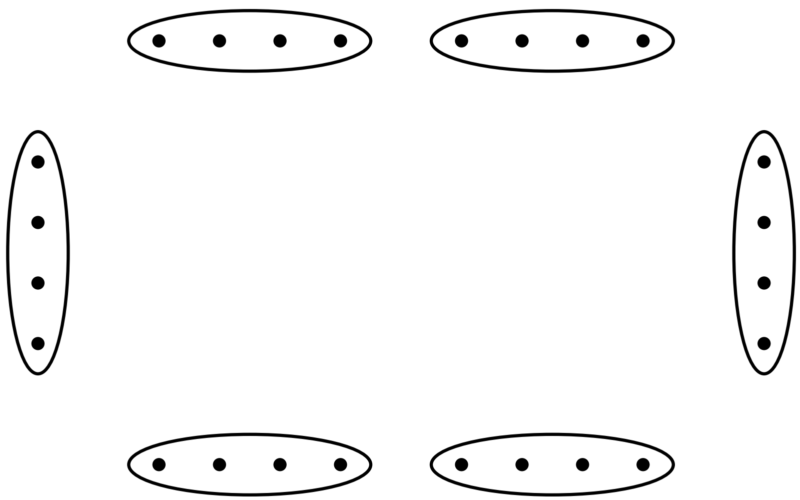
What about hypergraphs?

Let $\mathcal{G}_{n,r,s}$ denote a uniformly random r-regular s-uniform hypergraph on $[n]$. Here r, s are fixed constants. Assume $s|rn$.

What about hypergraphs?

Let $\mathcal{G}_{n,r,s}$ denote a uniformly random r-regular s-uniform hypergraph on $[n]$. Here r, s are fixed constants. Assume $s|rn$.

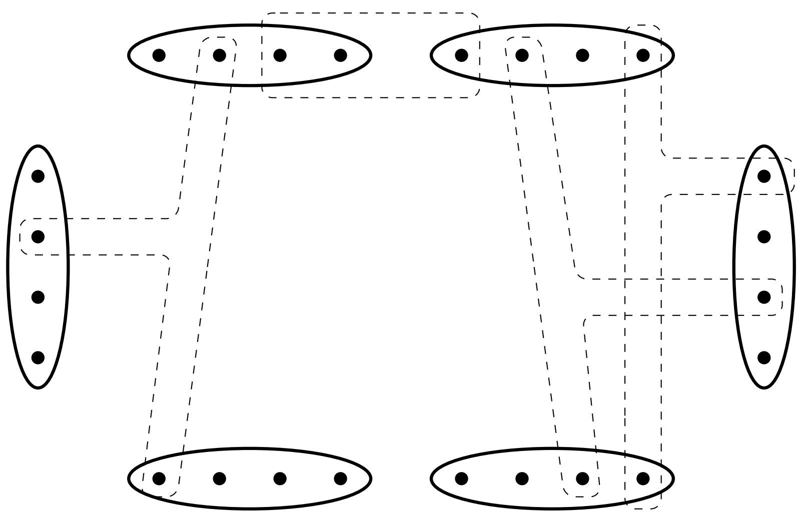
Calculations are performed in the configuration model.



What about hypergraphs?

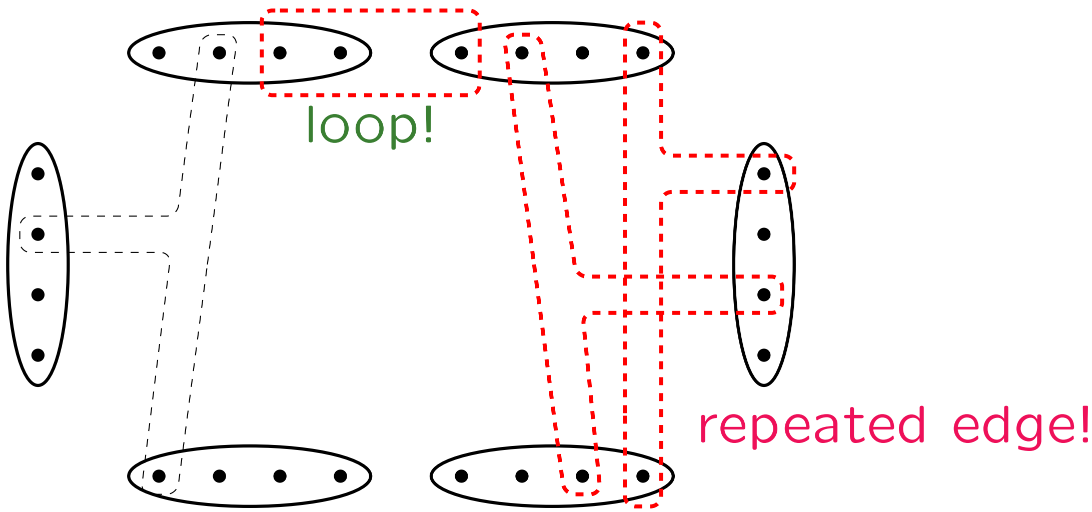
Let $\mathcal{G}_{n,r,s}$ denote a uniformly random r-regular s-uniform hypergraph on $[n]$. Here r, s are fixed constants. Assume $s|rn$.

Calculations are performed in the configuration model.



What about hypergraphs?
Let $\mathcal{G}_{n,r,s}$ denote a uniformly random r-regular s-uniform hypergraph on $[n]$. Here r, s are fixed constants. Assume $s|rn$.

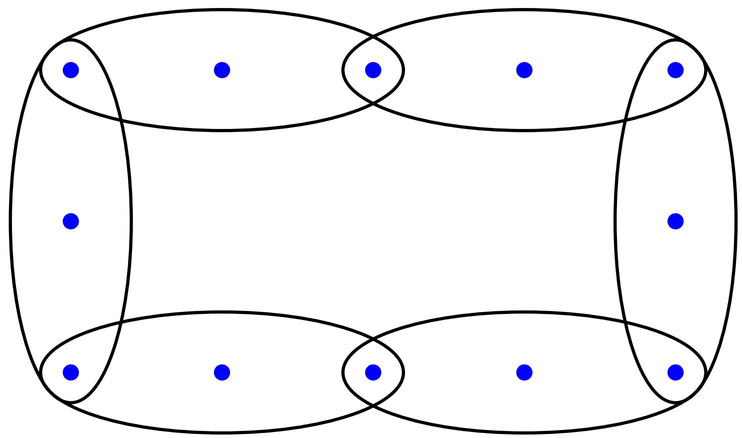
Calculations are performed in the configuration model.

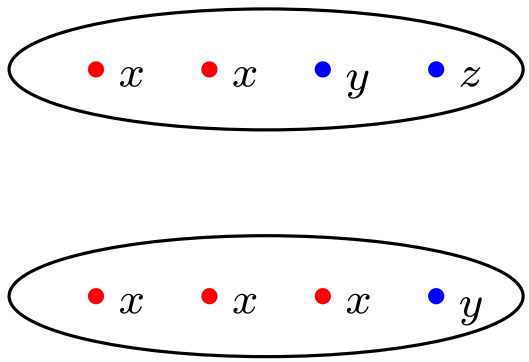


Cooper, Frieze, Molloy & Reed (1996): Pr(simple) $\sim e^{-\frac{(r-1)(s-1)}{2}}$.

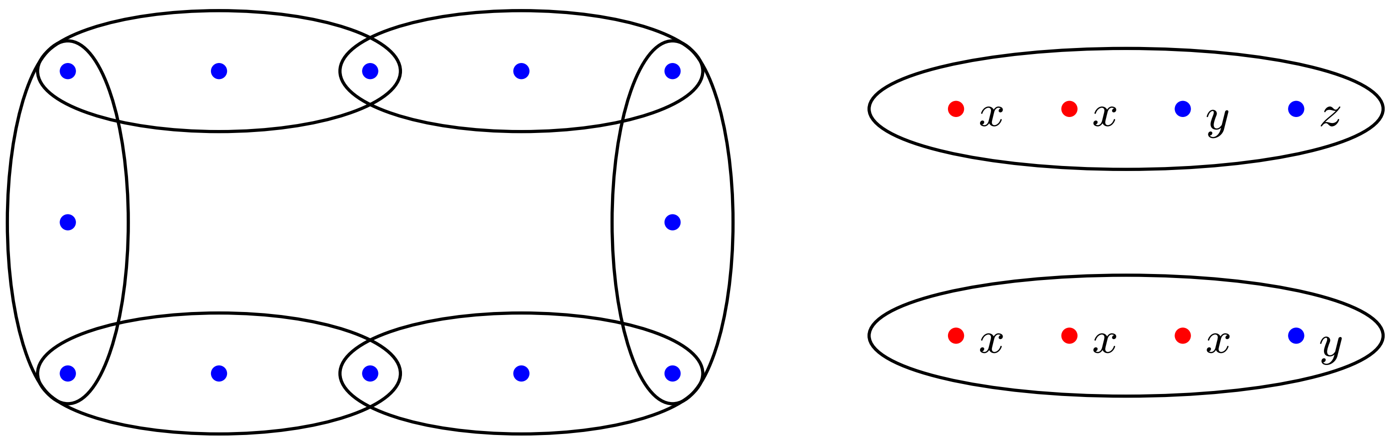
Let X_{in} be the number of loose i-cycles in $\mathcal{G}_{n,r,s}$, for $i \geq 2$, and let X_{1n} be the number of 1-cycles.

Let X_{in} be the number of loose i-cycles in $\mathcal{G}_{n,r,s}$, for $i \geq 2$, and let X_{1n} be the number of 1-cycles.





Let X_{in} be the number of loose i-cycles in $\mathcal{G}_{n,r,s}$, for $i \geq 2$, and let X_{1n} be the number of 1-cycles.



Cooper et al. (1996) proved that the X_{in} are asymptotically independent Poisson random variables, with

$$\mathbb{E}\, X_{in} \to \lambda_i = \frac{((r-1)(s-1))^i}{2i}.$$

So (A1) holds.

PROBLEM: In the configuration model, when $s \geq 3$, the event "is simple" is NOT captured by conditioning on the event $X_{1n} = X_{2n} = 0$, or on the event $X_{1n} = 0$.

PROBLEM: In the configuration model, when $s \geq 3$, the event "is simple" is NOT captured by conditioning on the event $X_{1n} = X_{2n} = 0$, or on the event $X_{1n} = 0$.

Since $\Pr(\text{repeated edge}) = o(1)$, conditional probabilities are no problem, but we must be careful with the expected value.

PROBLEM: In the configuration model, when $s \geq 3$, the event "is simple" is NOT captured by conditioning on the event $X_{1n} = X_{2n} = 0$, or on the event $X_{1n} = 0$.

Since $\Pr(\text{repeated edge}) = o(1)$, conditional probabilities are no problem, but we must be careful with the expected value.

$\Rightarrow$ The SSCM can't tell us $\frac{\mathbb{E}\, Y_{\mathcal{G}}}{\mathbb{E}\, Y}$.

Cooper, Frieze, Molloy & Reed (1996): existence threshold for perfect matchings, which a.a.s. exist in $\mathcal{G}_{n,r,s}$ when $s < \sigma_r$, where

$$\sigma_r = \frac{\log r}{(r-1)\log(r/(r-1))} + 1.$$

Cooper, Frieze, Molloy & Reed (1996): existence threshold for perfect matchings, which a.a.s. exist in $\mathcal{G}_{n,r,s}$ when $s < \sigma_r$, where

$$\sigma_r = \frac{\log r}{(r-1)\log(r/(r-1))} + 1.$$

Altman, Greenhill, Isaev, Ramadurai (2020):
existence threshold for loose Hamilton cycles, which a.a.s. exist in $\mathcal{G}_{n,r,s}$ when $r > \rho(s)$, where

$$\rho(s) \approx \frac{e^{s-1}}{s-1} - \frac{s-2}{2} + o_s(1).$$

(The $o_s(1)$ term tends to zero exponentially fast as $s \to \infty$.)

Greenhill, Isaev, Liang (arXiv:2005.07350):
existence threshold for spanning trees in $\mathcal{G}_{n,r,s}$.

Greenhill, Isaev, Liang (`arXiv:2005.07350`):
existence threshold for spanning trees in $\mathcal{G}_{n,r,s}$.

If $s \geq 5$ then spanning trees a.a.s. exist in $\mathcal{G}_{n,r,s}$ when $r > \rho(s)$, where

$$\rho(s) \approx \frac{e^{s-2}}{s-1} - \frac{s^2 - 3s + 1}{2(s-1)} + o_s(1).$$

Greenhill, Isaev, Liang (`arXiv:2005.07350`):
existence threshold for spanning trees in $\mathcal{G}_{n,r,s}$.

If $s \geq 5$ then spanning trees a.a.s. exist in $\mathcal{G}_{n,r,s}$ when $r > \rho(s)$, where

$$\rho(s) \approx \frac{e^{s-2}}{s-1} - \frac{s^2 - 3s + 1}{2(s-1)} + o_s(1).$$

If $s = 2, 3, 4$ then any $r \geq 2$ gives a.a.s. existence, except $(r, s) = (2, 2)$.

Greenhill, Isaev, Liang (`arXiv:2005.07350`):
existence threshold for spanning trees in $\mathcal{G}_{n,r,s}$.

If $s \geq 5$ then spanning trees a.a.s. exist in $\mathcal{G}_{n,r,s}$ when $r > \rho(s)$, where

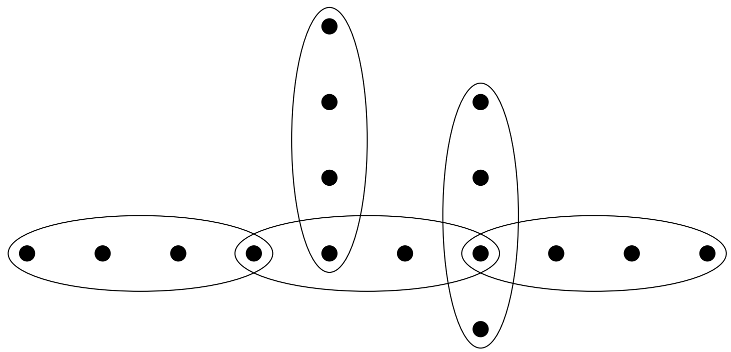
$$\rho(s) \approx \frac{e^{s-2}}{s-1} - \frac{s^2 - 3s + 1}{2(s-1)} + o_s(1).$$

If $s = 2, 3, 4$ then any $r \geq 3$ gives a.a.s. existence, except $(r, s) = (2, 2)$.

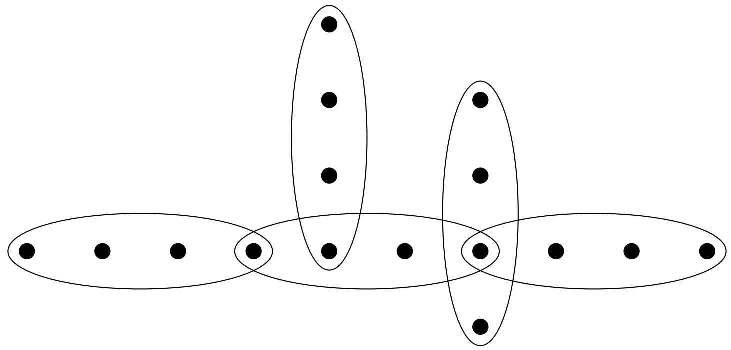
We build on earlier work by Greenhill, Kwan, Wind (2014) for graphs, which

- found expected number of spanning trees in $\mathcal{G}_{n,d}$ for $d \geq 3$,
- gave asymptotic distribution for cubic graphs.

A tree is connected and acyclic, where these terms are defined using Berge cycles and Berge paths. No 2-cycles means that edges overlap in at most 1 vertex (linear).



A tree is connected and acyclic, where these terms are defined using Berge cycles and Berge paths. No 2-cycles means that edges overlap in at most 1 vertex (linear).



A necessary condition for an s-uniform hypergraph on $[n]$ to contain a spanning tree is that

$$n = (s-1)t + 1$$

where $t = \frac{n-1}{s-1} \in \mathbb{Z}^{+}$ is the number of edges in the spanning tree.

Trees in uniform hypergraphs

Suppose that $n = (s-1)t + 1$ for some $t \in \mathbb{Z}^+$.

Bolian (1988) The number of s-uniform trees on $[n]$ is

$$n^{t-1}\,\frac{(n-1)!}{t!\,((s-1)!)^t}.$$

Trees in uniform hypergraphs

Suppose that $n = (s-1)t + 1$ for some $t \in \mathbb{Z}^+$.

Bolian (1988) The number of s-uniform trees on $[n]$ is

$$n^{t-1} \frac{(n-1)!}{t!\,((s-1)!)^t}.$$

When $s = 2$ we recover Cayley's formula (here $t = n - 1$).

Trees in uniform hypergraphs

Suppose that $n = (s-1)t + 1$ for some $t \in \mathbb{Z}^+$.

A tree degree sequence is a sequence $\boldsymbol{x} = (x_1, \ldots, x_n)$ of positive integers which sum to st.

Trees in uniform hypergraphs

Suppose that $n = (s-1)t + 1$ for some $t \in \mathbb{Z}^+$.

A tree degree sequence is a sequence $\boldsymbol{x} = (x_1, \ldots, x_n)$ of positive integers which sum to st.

Bacher (2011)
The number of s-uniform trees on $[n]$ with degree sequence $\boldsymbol{x}$ is

$$\frac{(s-1)\,(n-2)!}{((s-1)!)^t} \prod_{i=1}^{n} \frac{1}{(x_j - 1)!}.$$

Trees in uniform hypergraphs

Suppose that $n = (s-1)t + 1$ for some $t \in \mathbb{Z}^+$.

A tree degree sequence is a sequence $\boldsymbol{x} = (x_1, \ldots, x_n)$ of positive integers which sum to st.

Bacher (2011)
The number of s-uniform trees on $[n]$ with degree sequence $\boldsymbol{x}$ is

$$\frac{(s-1)\,(n-2)!}{((s-1)!)^t} \prod_{i=1}^{n} \frac{1}{(x_j - 1)!}.$$

This generalises the result of Moon (1970) in the graph case. These results can be proved using a hypergraph analogue of Prüfer codes.

Expected number

By summing over all tree degree sequences $\boldsymbol{x}$, we showed that the expected number of spanning trees in the configuration model is

$$\mathbb{E}\, Y = \frac{(s-1)(n-2)!}{((s-1)!)^t} \sum_{\boldsymbol{x}} \prod_{j=1}^{n} \frac{(r)_{x_j}}{(x_j-1)!} \frac{p(rn-st)}{p(rn)}$$

Expected number

By summing over all tree degree sequences $\boldsymbol{x}$, we showed that the expected number of spanning trees in the configuration model is

$$\mathbb{E}\,Y = \frac{(s-1)(n-2)!}{((s-1)!)^t} \sum_{\boldsymbol{x}} \prod_{j=1}^{n} \frac{(r)_{x_j}}{(x_j-1)!} \frac{p(rn-st)}{p(rn)}$$

$$= \frac{r^n\,(s-1)\,(n-2)!}{((s-1!)^t} \binom{(r-1)n}{t-1} \frac{p(rn-st)}{p(rn)}$$

where $p(sN)$ is the number of ways to partition sN points into N subsets (parts) of s points.

Apply Stirling's formula:

$$\mathbb{E}\,Y$$

$$\sim \frac{(r-1)^{1/2}(s-1)}{n(rs-r-s)^{\frac{s+1}{2(s-1)}}} \left(\frac{(s-1)^r\,(r-1)^{r-1}}{r^{(rs-r-s)}\,(rs-r-s)^{(rs-r-s)/(s-1)}} \right)^{n/s}.$$

Apply Stirling's formula:

$$\mathbb{E}\,Y$$

$$\sim \frac{(r-1)^{1/2}(s-1)}{n(rs-r-s)^{\frac{s+1}{2(s-1)}}} \left(\frac{(s-1)^r\,(r-1)^{r-1}}{r^{(rs-r-s)}\,(rs-r-s)^{(rs-r-s)/(s-1)}} \right)^{n/s}.$$

The behaviour is dominated by the base of the exponential: taking the logarithm, let

$$L_s(r) = \tfrac{r}{s}\log(s-1) + (r-1)\log(r-1) \\ - \tfrac{rs-r-s}{s}\log(r) - \tfrac{rs-r-s}{s(s-1)}\log(rs-r-s).$$

The behaviour is dominated by the base of the exponential: taking the logarithm, let

$$L_s(r) = \tfrac{r}{s}\log(s-1) + (r-1)\log(r-1) \\ \qquad - \tfrac{rs-r-s}{s}\log(r) - \tfrac{rs-r-s}{s(s-1)}\log(rs-r-s).$$

We proved that if $s \in \{2, 3, 4\}$ then $L_s(r) > 0$ for all $r \geq 2$, except $(r, s) = (2, 2)$.

The behaviour is dominated by the base of the exponential: taking the logarithm, let

$$L_s(r) = \tfrac{r}{s}\log(s-1) + (r-1)\log(r-1) \\ - \tfrac{rs-r-s}{s}\log(r) - \tfrac{rs-r-s}{s(s-1)}\log(rs-r-s).$$

We proved that if $s \in \{2,3,4\}$ then $L_s(r) > 0$ for all $r \geq 2$, except $(r,s) = (2,2)$.

For $s \geq 5$ there is a unique threshold $\rho(s) \in (2,\infty)$ so that

$$L_s(r) \text{ is } \begin{cases} < 0 & \text{for } r \in [2, \rho(s)), \\ > 0 & \text{for } r \in (\rho(s), \infty). \end{cases}$$

The behaviour is dominated by the base of the exponential: taking the logarithm, let

$$L_s(r) = \frac{r}{s}\log(s-1) + (r-1)\log(r-1) - \frac{rs-r-s}{s}\log(r) - \frac{rs-r-s}{s(s-1)}\log(rs-r-s).$$

We proved that if $s \in \{2,3,4\}$ then $L_s(r) > 0$ for all $r \geq 2$, except $(r,s) = (2,2)$.

For $s \geq 5$ there is a unique threshold $\rho(s) \in (2,\infty)$ so that

$$L_s(r) \quad \text{is} \quad \begin{cases} < 0 & \text{for } r \in [2,\rho(s)), \\ > 0 & \text{for } r \in (\rho(s),\infty). \end{cases}$$

s	5	6	7	8	9	10	11
$\rho(s)$	3.03	8.71	22.14	54.61	133.59	327.25	805.84

Short cycles

We calculated that in the configuration model,

$$\frac{\mathbb{E}(Y\,X_j)}{\mathbb{E}\,Y} \longrightarrow \lambda_j(1+\delta_j)$$

where

$$\delta_j = \frac{\left(\frac{r}{r-1} - s + 1\right)^j - 2}{((r-1)(s-1))^j}.$$

(Similar calculations for more than one cycle.)

Short cycles

We calculated that in the configuration model,

$$\frac{\mathbb{E}(YX_j)}{\mathbb{E}\,Y} \longrightarrow \lambda_j(1+\delta_j)$$

where

$$\delta_j = \frac{\left(\frac{r}{r-1} - s + 1\right)^j - 2}{((r-1)(s-1))^j}.$$

(Similar calculations for more than one cycle.)

Then we showed that (A2) and (A3) hold, and

$$\exp\left(\sum_{k=2}^{\infty} \lambda_k \delta_k^2\right) = \frac{r^2\sqrt{s-1}}{\sqrt{(r^2 - rs + r + s - 1)(rs - r - s)(r-1)}}.$$

Second moment

We must prove a certain 2-variable real function has a unique global maximum in the interior of a given bounded domain.

Second moment

We must prove a certain 2-variable real function has a unique global maximum in the interior of a given bounded domain.

We express the second moment as, up to a $(1 + o(1))$ factor,

$$\sum_{(k,b)\in D} \psi(k/n, b/n)\ \mathrm{exp}(n\,\varphi(k/n, b/n))$$

where k, b are two parameters arising from the combinatorics and D is the natural domain of these parameters. The function $\psi(\alpha, \beta)$ is relatively unimportant . . .

. . . and

$$\begin{aligned}\varphi(\alpha,\beta) = &(\alpha+\beta)\log(r-1) + g(\alpha+\beta) + g(r-1-\alpha-\beta)\\ &- \tfrac{2}{s-1}g(\beta) - g(\alpha) - \tfrac{1}{s(s-1)}g(rs-r-s-s\beta)\\ &- \tfrac{1}{s-1}g(1-(s-1)\alpha-\beta)\end{aligned}$$

where $g(x) = x\log x$ for $x > 0$, and $g(0) = 0$.

... and

$$\begin{aligned}\varphi(\alpha,\beta) = &(\alpha+\beta)\log(r-1) + g(\alpha+\beta) + g(r-1-\alpha-\beta)\\ &-\tfrac{2}{s-1}g(\beta) - g(\alpha) - \tfrac{1}{s(s-1)}g(rs-r-s-s\beta)\\ &-\tfrac{1}{s-1}g(1-(s-1)\alpha-\beta)\end{aligned}$$

where $g(x) = x\log x$ for $x > 0$, and $g(0) = 0$.

Lemma: Assume that $r, s \geq 2$ such that $r > \rho(s)$ when $s \geq 5$, or $r \geq 3$ when $s \in \{2, 3, 4\}$. Then φ has a unique maximum in the relevant domain at the point

$$\alpha_0 = \tfrac{1}{r(s-1)}, \qquad \beta_0 = \tfrac{rs-r-s}{r(s-1)}.$$

. . . and

$$\begin{aligned}\varphi(\alpha,\beta) =& (\alpha+\beta)\log(r-1) + g(\alpha+\beta) + g(r-1-\alpha-\beta) \\ & - \tfrac{2}{s-1} g(\beta) - g(\alpha) - \tfrac{1}{s(s-1)} g(rs-r-s-s\beta) \\ & - \tfrac{1}{s-1} g(1-(s-1)\alpha-\beta)\end{aligned}$$

where $g(x) = x\log x$ for $x > 0$, and $g(0) = 0$.

Lemma: Assume that $r, s \geq 2$ such that $r > \rho(s)$ when $s \geq 5$, or $r \geq 3$ when $s \in \{2,3,4\}$. Then φ has a unique maximum in the relevant domain at the point

$$\alpha_0 = \tfrac{1}{r(s-1)}, \qquad \beta_0 = \tfrac{rs-r-s}{r(s-1)}.$$

This implies that (A4) holds $\Rightarrow$ can apply SSCM to Y.

What about that PROBLEM going from $\mathbb{E}\,Y$ to $\mathbb{E}\,Y_{\mathcal{G}}$?

What about that PROBLEM going from $\mathbb{E}\,Y$ to $\mathbb{E}\,Y_{\mathcal{G}}$?

Happily, Aldosari & Greenhill (`arXiv:1907.04493`) used asymptotic enumeration, in a more general setting that covers constant r, s, to show that

$$\mathbb{E}\,Y_{\mathcal{G}} \sim e^{-\lambda_1 \delta_1}\,\mathbb{E}\,Y.$$

What about that PROBLEM going from $\mathbb{E}\, Y$ to $\mathbb{E}\, Y_{\mathcal{G}}$?

Happily, Aldosari & Greenhill (arXiv:1907.04493) used asymptotic enumeration, in a more general setting that covers constant r, s, to show that

$$\mathbb{E}\, Y_{\mathcal{G}} \sim e^{-\lambda_1 \delta_1}\, \mathbb{E}\, Y.$$

This leads to the existence threshold result, and gives us the asymptotic distribution: if $\mathbb{E}\, Y_{\mathcal{G}} \to \infty$ then

$$\frac{Y_{\mathcal{G}}}{\mathbb{E}\, Y_{\mathcal{G}}} \xrightarrow{d} \prod_{j=2}^{\infty} (1+\delta_j)^{Z_j}\, e^{\lambda_j \delta_j} \qquad \text{as} \quad n \to \infty.$$

Some ingredients in the proof

We used Generalised Jensen's identity: for $b \geq 2$,

$$\sum_{\substack{k_1+\cdots+k_b=m,\\ k_j\geq 0}} \prod_{i=1}^{b} \binom{x_i+ck_i}{k_i}$$

$$= \sum_{k=0}^{m} \binom{k+b-2}{k} \binom{x_1+\cdots+x_b+cm-k}{m-k} c^k.$$

Some ingredients in the proof

We used Generalised Jensen's identity: for $b \geq 2$,

$$\sum_{\substack{k_1+\cdots+k_b=m,\\ k_j\geq 0}} \prod_{i=1}^{b} \binom{x_i + ck_i}{k_i} = \sum_{k=0}^{m} \binom{k+b-2}{k} \binom{x_1 + \cdots + x_b + cm - k}{m-k} c^k.$$

This led to a more tractable form for the expression for the second moment, and enabled us to extend Greenhill, Kwan, Wind (2014).

Some ingredients in the proof

We used Generalised Jensen's identity: for $b \geq 2$,

$$\sum_{\substack{k_1+\cdots+k_b=m,\\ k_j\geq 0}} \prod_{i=1}^{b} \binom{x_i + ck_i}{k_i}$$

$$= \sum_{k=0}^{m} \binom{k+b-2}{k} \binom{x_1 + \cdots + x_b + cm - k}{m-k} c^k.$$

This led to a more tractable form for the expression for the second moment, and enabled us to extend Greenhill, Kwan, Wind (2014).

Also generating functions (for short cycles) and a Laplace summation theorem from Greenhill, Janson and Ruciński (2010) to help with the second moment calculations.

Greenhill, Janson, Ruciński (2010), Laplace summation tool.

Say you want to evaluate

$$\sum_{\boldsymbol{\ell} \in (\mathcal{L} + \boldsymbol{\ell}_n) \cap nK} a_n(\ell)$$

where

Greenhill, Janson, Ruciński (2010), Laplace summation tool.

Say you want to evaluate

$$\sum_{\boldsymbol{\ell} \in (\mathcal{L} + \boldsymbol{\ell}_n) \cap nK} a_n(\boldsymbol{\ell})$$

where

$\mathcal{L} \subseteq \mathbb{R}^m$ is a lattice with full rank,

$\boldsymbol{\ell}_n$ is a shift vector,

Greenhill, Janson, Ruciński (2010), Laplace summation tool.

Say you want to evaluate

$$\sum_{\boldsymbol{\ell} \in (\mathcal{L} + \boldsymbol{\ell}_n) \cap nK} a_n(\boldsymbol{\ell})$$

where

$\mathcal{L} \subseteq \mathbb{R}^m$ is a lattice with full rank,

$\boldsymbol{\ell}_n$ is a shift vector,

$K \subset \mathbb{R}^m$ is a compact convex set with non-empty interior,

$a_n(\boldsymbol{\ell})$ is a product of factorials and powers.

If (away from the boundary)

$$a_n(\boldsymbol{\ell}) \sim b_n\, \psi(\boldsymbol{\ell}/n)\ \exp(n\varphi(\boldsymbol{\ell}/n))$$

If (away from the boundary)

$$a_n(\boldsymbol{\ell}) \sim b_n \, \psi(\boldsymbol{\ell}/n) \, \exp(n\varphi(\boldsymbol{\ell}/n))$$

and $\varphi(\boldsymbol{x})$ has a unique maximum in the interior of K, at $\boldsymbol{x}_0$ and a couple of other mild conditions,

If (away from the boundary)

$$a_n(\boldsymbol{\ell}) \sim b_n\, \psi(\boldsymbol{\ell}/n)\ \exp(n\varphi(\boldsymbol{\ell}/n))$$

and $\varphi(\boldsymbol{x})$ has a unique maximum in the interior of K, at $\boldsymbol{x}_0$ and a couple of other mild conditions, then

$$\sum_{\ell \in (\mathcal{L}+\boldsymbol{\ell}_n)\cap nK} a_n(\ell) \sim \frac{b_n\,(2\pi n)^{m/2}\,\psi(\boldsymbol{x}_0)}{\det(\mathcal{L})\ \det(-H_0)^{-1/2}}\ \exp\left(n\varphi(\boldsymbol{x}_0)\right)$$

If (away from the boundary)

$$a_n(\boldsymbol{\ell}) \sim b_n \, \psi(\boldsymbol{\ell}/n) \; \exp(n\varphi(\boldsymbol{\ell}/n))$$

and $\varphi(\boldsymbol{x})$ has a unique maximum in the interior of K, at $\boldsymbol{x}_0$ and a couple of other mild conditions, then

$$\sum_{\boldsymbol{\ell} \in (\mathcal{L} + \boldsymbol{\ell}_n) \cap nK} a_n(\ell) \sim \frac{b_n \, (2\pi n)^{m/2} \, \psi(\boldsymbol{x}_0)}{\det(\mathcal{L}) \; \det(-H_0)^{-1/2}} \exp\left(n\varphi(\boldsymbol{x}_0)\right)$$

where

$\det(\mathcal{L})$ is the determinant of the lattice $\mathcal{L}$,

and H_0 is the Hessian of φ at $\boldsymbol{x}_0$.