Graphon limit and large independent sets in uniform random cographs

Valentin Féray joint work with F. Bassino, M. Bouvel, M. Drmota, L. Gerin, M. Maazoun and A. Pierrot

CNRS, Institut Élie Cartan de Lorraine (IECL)

Workshop Graph Limits, Non-Parametric Models, and Estimation Berkeley, September 2022







V. Féray (CNRS, IECL)

Definition

A cograph is a P_4 -free graph, i.e. not containing P_4 as an induced subgraph.



Questions: asymptotic behaviour of a uniform random cograph

- What is its graphon limit?
- Existence of independent sets/cliques of linear size?

Motivations:

- When studying *H*-free graphs, the case *H* = *P*₄ is special with an interesting limit object;
- Probabilistic work around Erdős-Hajnal conjecture.

First part

Asymptotic enumeration of *H*-free graphs

Fix a graph *H*, and consider its chromatic number $k = \chi(H)$. Observation: if $\chi(G) \le k-1$, then *G* is *H* free.

Fix a graph H, and consider its chromatic number $k = \chi(H)$. Observation: if $\chi(G) \leq k-1$, then G is H free.

Easy: There are $\geq 2^{\left(1-\frac{1}{k-1}+o(1)\right)\binom{n}{2}}$ graphs *G* with *n* vertices and $\chi(G) \leq k-1$



Fix a graph *H*, and consider its chromatic number $k = \chi(H)$. Observation: if $\chi(G) \le k-1$, then *G* is *H* free.

Easy: There are
$$\ge 2^{\left(1-\frac{1}{k-1}+o(1)\right)\binom{n}{2}}$$
 graphs *G* with *n* vertices and $\chi(G) \le k-1$

Conclusion:

$$\left| \left\{ \begin{array}{c} H \text{-free graphs} \\ \text{with } n \text{ vertices} \end{array} \right\} \right| \geq 2^{\left(1 - \frac{1}{k-1} + o(1)\right)\binom{n}{2}}$$

Fix a graph *H*, and consider its chromatic number $k = \chi(H)$. Observation: if $\chi(G) \le k-1$, then *G* is *H* free.

Easy: There are
$$\ge 2^{\left(1-\frac{1}{k-1}+o(1)\right)\binom{n}{2}}$$
 graphs *G* with *n* vertices and $\chi(G) \le k-1$

Conclusion:

$$\left| \left\{ \begin{array}{c} H \text{-free graphs} \\ \text{with } n \text{ vertices} \end{array} \right\} \right| \geq 2^{\left(1 - \frac{1}{k-1} + o(1)\right)\binom{n}{2}}$$

 \rightarrow works also if $\chi(\overline{H}) = k$ (putting cliques in the blue clusters),

Fix a graph *H*, and consider its chromatic number $k = \chi(H)$. Observation: if $\chi(G) \le k-1$, then *G* is *H* free.

Easy: There are
$$\ge 2^{\left(1-\frac{1}{k-1}+o(1)\right)\binom{n}{2}}$$
 graphs *G* with *n* vertices and $\chi(G) \le k-1$

Conclusion:

$$\left| \left\{ \begin{array}{l} H \text{-free graphs} \\ \text{with } n \text{ vertices} \end{array} \right\} \right| \geq 2^{\left(1 - \frac{1}{k-1} + o(1)\right)\binom{n}{2}}$$

 \rightarrow works also if $\chi(\overline{H}) = k$ (putting cliques in the blue clusters), or if V_H cannot be partitioned in *s* cliques and *t* independent sets for some given s + t = k - 1.

Theorem (Prömel-Steger, '92)

$$\left\{ \begin{array}{c} H\text{-free graphs} \\ with n vertices \end{array} \right\} = 2^{\left(1 - \frac{1}{r} + o(1)\right)\binom{n}{2}},$$

where r is maximal such that V_H cannot be partitioned into s cliques and t independent sets, for some (s, t) with s + t = r.

r+1: coloring number of H.

 $H = P_4$ is one of the few cases with r = 1. Hence

$$c_n := \left| \left\{ \begin{array}{c} \text{cographs} \\ \text{with } n \text{ vertices} \end{array} \right\} \right| = 2^{o(n^2)}.$$

Background: enumerating cographs

Operations on graphs:



disjoint union $G_1 \cup G_2$

join $G_1 + G_2$

 G_2

Proposition (Corneil-Lerchs-Stewart Burlingham '81)

The class of cographs is the smallest set of graphs containing the one-vertex graph, and stable by disjoint unions and joins (cographs are sometimes called complement-reducible graphs).

Background: enumerating cographs

Operations on graphs:





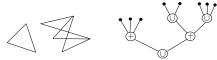
join $G_1 + G_2$

 G_2

Proposition (Corneil-Lerchs-Stewart Burlingham '81)

The class of cographs is the smallest set of graphs containing the one-vertex graph, and stable by disjoint unions and joins (cographs are sometimes called complement-reducible graphs).

Consequence: cographs can be encoded by decorated trees



A cograph and the associated decorated tree

Easy to enumerate: $c_n \sim C n! \kappa^n n^{-3/2}$, with $\kappa = (2\log(2) - 1)^{-1}$ and some explicit constant C (labeled case).

V. Féray (CNRS, IECL)

Random cographs

Second part

Graphon limit of uniform random cographs

A general result for H-free graphs

Fix H and let G_n be a uniform H-free graph on n vertices.

Question: what is the graphon limit of G_n (if it exists)?

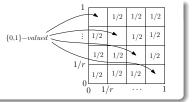
A general result for H-free graphs

Fix H and let G_n be a uniform H-free graph on n vertices.

Question: what is the graphon limit of G_n (if it exists)?

Theorem (Hatami, Janson, Szegedy, '18)

Let H be a graph and take r as before. Then any subsequential limit of G_n is supported on the set of graphons of the following form:



Reminiscent of the picture



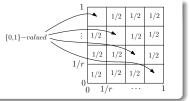
A general result for H-free graphs

Fix H and let G_n be a uniform H-free graph on n vertices.

Question: what is the graphon limit of G_n (if it exists)?

Theorem (Hatami, Janson, Szegedy, '18)

Let H be a graph and take r as before. Then any subsequential limit of G_n is supported on the set of graphons of the following form:



It does not say much for $H = P_4$ (where r = 1). In fact, all P_4 -free graphons are $\{0, 1\}$ -valued.

Limit of uniform random cograph

Theorem (Bouvel-Bassino-F.-Gerin-Maazoun-Pierrot '22, Stufler '22)

Let G_n be a uniform random (either labeled or unlabeled) cograph with n vertices. Then W_{G_n} converges in distribution to a random graphon W^{Br} , which we call Brownian cographon.

Limit of uniform random cograph

Theorem (Bouvel-Bassino-F.-Gerin-Maazoun-Pierrot '22, Stufler '22)

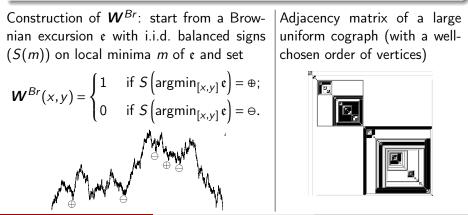
Let G_n be a uniform random (either labeled or unlabeled) cograph with n vertices. Then W_{G_n} converges in distribution to a random graphon W^{Br} , which we call Brownian cographon.

Construction of
$$W^{Br}$$
: start from a Brow-
nian excursion \mathfrak{e} with i.i.d. balanced signs
 $(S(m))$ on local minima m of \mathfrak{e} and set
 $W^{Br}(x,y) = \begin{cases} 1 & \text{if } S\left(\operatorname{argmin}_{[x,y]}\mathfrak{e}\right) = \oplus; \\ 0 & \text{if } S\left(\operatorname{argmin}_{[x,y]}\mathfrak{e}\right) = \Theta. \end{cases}$

V. Féray (CNRS, IECL)

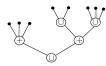
Limit of uniform random cograph

Theorem (Bouvel–Bassino–F.–Gerin–Maazoun–Pierrot '22, Stufler '22) Let G_n be a uniform random (either labeled or unlabeled) cograph with n vertices. Then W_{G_n} converges in distribution to a random graphon W^{Br} , which we call Brownian cographon.



Heuristic for the theorem

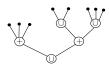
• A cograph *G* is encoded by a decorated tree *T*;



- Vertices in *G* correspond to leaves in *T*;
- Vertices v₁ and v₂ are connected if their youngest common ancestor is decorated by ⊕;

Heuristic for the theorem

• A cograph *G* is encoded by a decorated tree *T*;



- Vertices in *G* correspond to leaves in *T*;
- Vertices v₁ and v₂ are connected if their youngest common ancestor is decorated by ⊕;

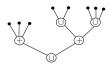
• The limit of T is Aldous' Continuum Random Tree T_{∞} , coded by a Brownian excursion e;



- Leaves of T_∞ form a measure 1 subset of [0, 1];
- Youngest common ancestor between x and y correspond to argmin_[x,y] e. Thus x, y are linked in W^{Br} if S(argmin_[x,y] e) = ⊕.

Heuristic for the theorem

• A cograph *G* is encoded by a decorated tree *T*;



- Vertices in *G* correspond to leaves in *T*;
- Vertices v₁ and v₂ are connected if their youngest common ancestor is decorated by ⊕;

• The limit of T is Aldous' Continuum Random Tree T_{∞} , coded by a Brownian excursion e;



- Leaves of T_{∞} form a measure 1 subset of [0, 1];
- Youngest common ancestor between x and y correspond to argmin_[x,y] e. Thus x, y are linked in W^{Br} if S(argmin_[x,y] e) = ⊕.

Note: in the discrete, decorations alternate; in the continuous, they are independent.

V. Féray (CNRS, IECL)

Random cographs

Sampling from W^{Br}

 $G(n, \boldsymbol{W}^{Br})$ has vertex set [n] and $i \sim j$ if and only if $\boldsymbol{W}^{Br}(U_i, U_j) = 1$ $(U_1, \dots, U_n \text{ i.i.d. unif. in } [0, 1]).$

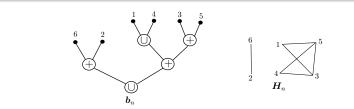
Sampling from W^{Br}

 $G(n, \boldsymbol{W}^{Br})$ has vertex set [n] and $i \sim j$ if and only if $\boldsymbol{W}^{Br}(U_i, U_j) = 1$ $(U_1, \dots, U_n \text{ i.i.d. unif. in } [0, 1]).$

Proposition

Let T_n be a uniform random binary tree with n leaves and choose independent {U, +} decoration for its internal node. Let H_n be the associated cograph. Then

$$\boldsymbol{G}(n, \boldsymbol{W}^{Br}) \stackrel{d}{=} \boldsymbol{H}_n.$$



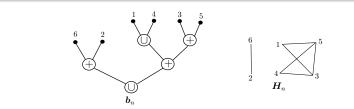
Sampling from W^{Br}

 $G(n, \boldsymbol{W}^{Br})$ has vertex set [n] and $i \sim j$ if and only if $\boldsymbol{W}^{Br}(U_i, U_j) = 1$ $(U_1, \dots, U_n \text{ i.i.d. unif. in } [0, 1]).$

Proposition

Let T_n be a uniform random binary tree with n leaves and choose independent {U, +} decoration for its internal node. Let H_n be the associated cograph. Then

$$\boldsymbol{G}(n, \boldsymbol{W}^{Br}) \stackrel{d}{=} \boldsymbol{H}_n.$$



Expected degree distribution of W^{Br}

Consider $G(n, W^{Br})$ the random graph with *n* vertices sampled from W^{Br} .

Proposition (Bouvel-Bassino-F.-Gerin-Maazoun-Pierrot '22)

The degree of vertex 1 in $G(n, W^{Br})$ is a random variable following the uniform distribution in $\{0, 1, ..., n-1\}$.

Expected degree distribution of W^{Br}

Consider $G(n, W^{Br})$ the random graph with *n* vertices sampled from W^{Br} .

Proposition (Bouvel-Bassino-F.-Gerin-Maazoun-Pierrot '22)

The degree of vertex 1 in $G(n, W^{Br})$ is a random variable following the uniform distribution in $\{0, 1, ..., n-1\}$.

Corollary 1

The expected degree distribution of \boldsymbol{W}^{Br}

$$\mathsf{Law}\left(\mathbb{E}\left[\int_{0}^{1}\boldsymbol{W}^{Br}(\boldsymbol{U},\boldsymbol{y})d\boldsymbol{y}\middle|\boldsymbol{U}\right]\right)$$

is the uniform distribution on [0,1].

Expected degree distribution of W^{Br}

Consider $G(n, W^{Br})$ the random graph with *n* vertices sampled from W^{Br} .

Proposition (Bouvel-Bassino-F.-Gerin-Maazoun-Pierrot '22)

The degree of vertex 1 in $G(n, W^{Br})$ is a random variable following the uniform distribution in $\{0, 1, ..., n-1\}$.

Corollary 1

The expected degree distribution of \boldsymbol{W}^{Br}

$$\mathsf{Law}\left(\mathbb{E}\left[\int_{0}^{1}\boldsymbol{W}^{Br}(\boldsymbol{U},\boldsymbol{y})d\boldsymbol{y}\middle|\boldsymbol{U}\right]\right)$$

is the uniform distribution on [0,1].

Corollary 2

The normalized degree $\frac{d_{\mathbf{v}}}{n}$ of a uniform random vertex \mathbf{v} in a uniform random cograph on n vertices is asymptotically uniform in [0,1].

V. Féray (CNRS, IECL)

Third part

Independent sets in uniform random cographs

Erdős-Hajnal conjecture ('89)

Fix a graph *H*. There exists $\varepsilon = \varepsilon(H)$ such that every *H*-free graph contains a homogeneous set of size n^{ε} .

homogeneous set = clique or independent set

Erdős-Hajnal conjecture ('89)

Fix a graph *H*. There exists $\varepsilon = \varepsilon(H)$ such that every *H*-free graph contains a homogeneous set of size n^{ε} .

Theorem (Loebl–Reed–Scott–Thomason–Thomassé, '14)

Fix a graph H. There exists $\varepsilon = \varepsilon(H)$ such that a uniform random H-free graph contains a homogeneous set of size n^{ε} (with high probability).

Erdős-Hajnal conjecture ('89)

Fix a graph *H*. There exists $\varepsilon = \varepsilon(H)$ such that every *H*-free graph contains a homogeneous set of size n^{ε} .

Theorem (Loebl-Reed-Scott-Thomason-Thomassé, '14)

Fix a graph H. There exists $\varepsilon = \varepsilon(H)$ such that a uniform random H-free graph contains a homogeneous set of size n^{ε} (with high probability).

Theorem (Kang-McDiarmid-Reed-Scott, '14)

For most H, the exists b = b(H) such that a uniform random H-free graph contains a homogeneous set of size bn (with high probability).

Erdős-Hajnal conjecture ('89)

Fix a graph *H*. There exists $\varepsilon = \varepsilon(H)$ such that every *H*-free graph contains a homogeneous set of size n^{ε} .

Theorem (Loebl-Reed-Scott-Thomason-Thomassé, '14)

Fix a graph H. There exists $\varepsilon = \varepsilon(H)$ such that a uniform random H-free graph contains a homogeneous set of size n^{ε} (with high probability).

Theorem (Kang-McDiarmid-Reed-Scott, '14)

For most H, the exists b = b(H) such that a uniform random H-free graph contains a homogeneous set of size bn (with high probability).

Question (KMRS, '14)

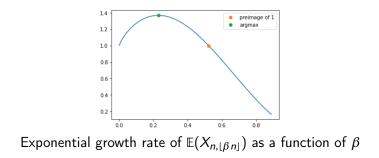
Does this hold for $H = P_4$?

Large independent sets in uniform random cographs

Theorem (Bouvel-Bassino-Drmota-F.-Gerin-Maazoun-Pierrot '22)

Let G_n be a uniform random cograph of size n.

• There exists $\beta_0 > 0$ s.t. for any $\beta < \beta_0$, the expected number $\mathbb{E}(X_{n,\lfloor\beta\,n\rfloor})$ of independent sets of size $\lfloor\beta\,n\rfloor$ in G_n grows exponentially fast.



Large independent sets in uniform random cographs

Theorem (Bouvel-Bassino-Drmota-F.-Gerin-Maazoun-Pierrot '22)

Let G_n be a uniform random cograph of size n.

- There exists β₀ > 0 s.t. for any β < β₀, the expected number E(X_{n,\βn]}) of independent sets of size [βn] in G_n grows exponentially fast.
- 2 The size of the largest independent set in G_n is $o_P(n)$.

From ②, for any $\beta > 0$, we have $X_{n,\lfloor\beta n\rfloor} = 0$ with high probability (the expectation is misleading).

Large independent sets in uniform random cographs

Theorem (Bouvel-Bassino-Drmota-F.-Gerin-Maazoun-Pierrot '22)

Let G_n be a uniform random cograph of size n.

There exists β₀ > 0 s.t. for any β < β₀, the expected number E(X_{n,\βn]}) of independent sets of size [βn] in G_n grows exponentially fast.

2 The size of the largest independent set in G_n is $o_P(n)$.

From ②, for any $\beta > 0$, we have $X_{n,\lfloor\beta n\rfloor} = 0$ with high probability (the expectation is misleading).

- ② answers KMRS's question from previous slide by the negative;
- Proof of 1 uses analytic combinatorics (and the tree encoding);
- Proof of 2 uses the graphon limit (next few slides).

Independence number of a graphon

Definition (Hladkỳ and Rocha, '20)

An independent set I of a graphon W is a subset $I \subseteq [0,1]$ such that W(x,y) = 0 for almost all (x,y) in $I \times I$. The independence number of W, denoted $\alpha(W)$, is the maximum measure of an independent set of W.

Clearly, if G is a graph with n vertices, then $n\alpha(W_G)$ is the maximum size of an independent set of G.

Independence number of a graphon

Definition (Hladkỳ and Rocha, '20)

An independent set I of a graphon W is a subset $I \subseteq [0,1]$ such that W(x,y) = 0 for almost all (x,y) in $I \times I$. The independence number of W, denoted $\alpha(W)$, is the maximum measure of an independent set of W.

Clearly, if G is a graph with n vertices, then $n\alpha(W_G)$ is the maximum size of an independent set of G.

Proposition (Hladkỳ and Rocha, '20) α is a lower semi-continuous function, i.e. if W_n converges to W, then $\limsup \alpha(W_n) \le \alpha(W)$.

We only need to prove:

Proposition

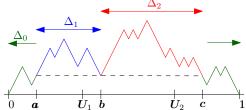
 $\alpha(\boldsymbol{W}^{Br}) = 0 \ a.s.$

Indeed, let G_n be a uniform random cograph on *n* vertices. Since W_{G_n} tends to W^{Br} and α is lower semi-continuous, it would imply $W_{G_n} \rightarrow 0$.

Proposition

 $\alpha(\mathbf{W}^{Br}) = 0 \ a.s.$

Sketch of proof: we use Aldous' self-similarity property of the Brownian excursion

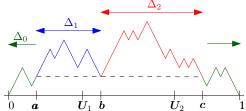


Starting from \mathfrak{e} , U_1 , U_2 , we get three independent excursions $(\mathfrak{e}_0, \mathfrak{e}_1, \mathfrak{e}_2)$ scaled by random lengths $(\Delta_0, \Delta_1, \Delta_2)$.

Proposition

 $\alpha(\mathbf{W}^{Br}) = 0 \ a.s.$

Sketch of proof: we use Aldous' self-similarity property of the Brownian excursion



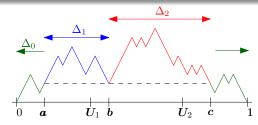
Starting from \mathfrak{e} , U_1 , U_2 , we get three independent excursions ($\mathfrak{e}_0, \mathfrak{e}_1, \mathfrak{e}_2$) scaled by random lengths ($\Delta_0, \Delta_1, \Delta_2$).

Starting from W^{Br} , we get three independent copies W_0^{Br} , W_1^{Br} and W_2^{Br} of the Brownian cographon.

V. Féray (CNRS, IECL)

Proposition

 $\alpha(\mathbf{W}^{Br}) = 0 \ a.s.$



An independent set of W^{Br} consists of

- an independent set of W_0^{Br} (scaled by Δ_0);
- if S(b) = Θ, an independent set of W₁^{Br} (scaled by Δ₁) and an independent set of W₂^{Br} (scaled by Δ₂);
- if S(b) = ⊕, an independent set of W₁^{Br} (scaled by Δ₁) or an independent set of W₂^{Br} (scaled by Δ₂);

Proposition

 $\alpha(\mathbf{W}^{Br}) = 0 \ a.s.$

Therefore, we have:

$$\alpha(W^{Br}) \leq \Delta_0 \alpha(W_0^{Br}) + \begin{cases} \Delta_1 \alpha(W_1^{Br}) + \Delta_2 \alpha(W_2^{Br}) & \text{if } S(b) = \Theta; \\ \max\left(\Delta_1 \alpha(W_1^{Br}), \Delta_2 \alpha(W_2^{Br})\right) & \text{if } S(b) = \Theta. \end{cases}$$

Proposition

 $\alpha(\mathbf{W}^{Br}) = 0 \ a.s.$

Therefore, we have:

$$\alpha(W^{Br}) \leq \Delta_0 \alpha(W_0^{Br}) + \begin{cases} \Delta_1 \alpha(W_1^{Br}) + \Delta_2 \alpha(W_2^{Br}) & \text{if } S(b) = \Theta; \\ \max\left(\Delta_1 \alpha(W_1^{Br}), \Delta_2 \alpha(W_2^{Br})\right) & \text{if } S(b) = \Theta. \end{cases}$$

 $\alpha(W_0^{Br})$, $\alpha(W_1^{Br})$ and $\alpha(W_2^{Br})$ are independent copies of $\alpha(W^{Br})$. This is an almost sure inequality, which implies an inequality satisfied by the law of $\alpha(W_0^{Br})$.

Proposition

 $\alpha(\mathbf{W}^{Br}) = 0 \ a.s.$

Therefore, we have:

$$\alpha(W^{Br}) \leq \Delta_0 \alpha(W_0^{Br}) + \begin{cases} \Delta_1 \alpha(W_1^{Br}) + \Delta_2 \alpha(W_2^{Br}) & \text{if } S(b) = \Theta; \\ \max\left(\Delta_1 \alpha(W_1^{Br}), \Delta_2 \alpha(W_2^{Br})\right) & \text{if } S(b) = \Theta. \end{cases}$$

 $\alpha(W_0^{Br})$, $\alpha(W_1^{Br})$ and $\alpha(W_2^{Br})$ are independent copies of $\alpha(W^{Br})$. This is an almost sure inequality, which implies an inequality satisfied by the law of $\alpha(W_0^{Br})$.

We show by a fixed point + monotonicity argument that δ_0 is the only distribution satisfying this inequality. Thus $\alpha(\mathbf{W}^{Br}) = 0$ a.s.

⁽⁹⁾ We do not control the speed of convergence of $\alpha(W_{G_n})$ to 0.

Summary

	enumeration	graphon limit	independent set
<i>H</i> -free (<i>r</i> > 1)	$e^{\Theta(n^2)}$	$\{0,1\}-valued \begin{array}{c c c c c c c c c c c c c c c c c c c $	$\Theta_{\mathcal{P}}(n)$ (for most H)
cographs	$e^{n\log(n)+\Theta(n)}$ (labeled)	W ^{Br}	0 _P (n)

Summary

	enumeration	graphon limit	independent set
<i>H</i> -free (<i>r</i> > 1)	$e^{\Theta(n^2)}$	$\{0.1\}-valued \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\Theta_P(n)$ (for most H)
cographs	$e^{n\log(n)+\Theta(n)}$ (labeled)	W ^{Br}	0 _P (n)

Thank you for your attention

V. Féray (CNRS, IECL)