Graphon limits of static and dynamic models of random cographs

Valentin Féray
joint work with F. Bassino, M. Bouvel, L. Gerin, M. Maazoun, A. Pierrot and K. Rivera-Lopez

CNRS, Institut Élie Cartan de Lorraine (IECL)

Virtual Informal Systems Seminar, online, May 12th, 2023

What is this talk about?

Cographs: nice class of graphs (definition on next slide), well-understood from a combinatorial/algorithmic point of view.

Here: a probabilistic/large network perspective on cographs. In particular, we will describe *graphon limits* of three models of random cographs.

What is this talk about?

Cographs: nice class of graphs (definition on next slide), well-understood from a combinatorial/algorithmic point of view.

Here: a probabilistic/large network perspective on cographs. In particular, we will describe *graphon limits* of three models of random cographs.

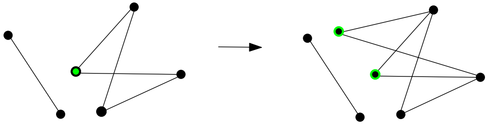
Motivations:

- Simple models which have some nontrivial graphon limit and limiting dynamics in the space of graphons.
- Probabilistic work around Erdős-Hajnal conjecture.

Cographs (1/2)

Let G be a graph. A duplication operation consist in

- choosing a vertex of G;
- adding a new vertex v' with the same neighbours as v;
- possibly connect v and v'.



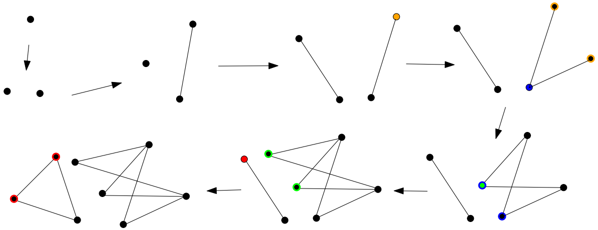
Cographs (1/2)

Let G be a graph. A duplication operation consist in

- choosing a vertex of G;
- adding a new vertex v' with the same neighbours as v;
- possibly connect v and v'.

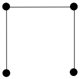
Definition

A cograph is a graph that can be obtained starting from the one-vertex graph and iterating duplication operations.



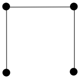
Cographs (2/2)

Observation: the path P_4 is not a cograph.



Cographs (2/2)

Observation: the path P_4 is not a cograph.

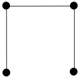


Proposition (direct characterization of cographs, Corneil–Lerchs–Stewart Burlingham '81)

A graph is a cograph if it one **cannot** *find four (distinct) vertices* v_1, v_2, v_3, v_4 *of* G *such that the induced graph* $G[v_1, v_2, v_3, v_4]$ *is* P_4.

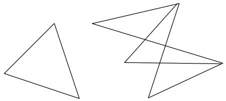
Cographs (2/2)

Observation: the path P_4 is not a cograph.

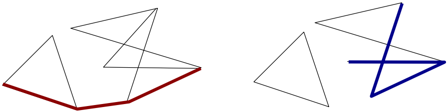


Proposition (direct characterization of cographs, Corneil–Lerchs–Stewart Burlingham '81)

A graph is a cograph if it one **cannot** *find four (distinct) vertices v_1, v_2, v_3, v_4 of G such that the induced graph $G[v_1, v_2, v_3, v_4]$ is P_4.*



A (disconnected) cograph



These are not cographs

Cographs are well-studied: many other characterizations, recognition algorithms, tree decomposition, . . .

Models of random cographs

1. G_n: uniform random cograph with n vertices;
2. H_n^p: constructed recursively by duplicating a uniform random vertex in H_{n-1}^p and connecting the two new vertices with probability p in $[0,1]$ (starting with $H_1^p = \bullet$).
3. $X_n^p(t)$: obtained by duplicating a uniform random vertex (connecting the two new vertices with proba p) and then deleting a uniform random vertex in $X_n^p(t-1)$, starting with *any* graph $X_n^p(0)$ with n vertices.

Models of random cographs

1. G_n: uniform random cograph with n vertices;
2. H_n^p: constructed recursively by duplicating a uniform random vertex in H_{n-1}^p and connecting the two new vertices with probability p in $[0,1]$ (starting with $H_1^p = \bullet$).
3. $X_n^p(t)$: obtained by duplicating a uniform random vertex (connecting the two new vertices with proba p) and then deleting a uniform random vertex in $X_n^p(t-1)$, starting with *any* graph $X_n^p(0)$ with n vertices.

Motivations:

1. random graph theory (probabilistic work around Erdős-Hajnal conjecture) ;

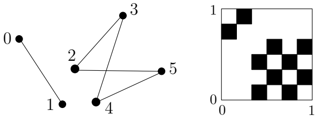
3. similar model with other combinatorial objects (partitions) coming from population dynamics ;

2. appears in the study of 3.

A short course on graphons (1/2)

Graph function: with a graph G on vertex-set $\{0,\dots,n-1\}$, we associate its *rescaled adjacency matrix/pixel picture* $W_G : [0,1]^2 \to [0,1]$

$$W_G(x,y) = \begin{cases} 1 & \text{if } \{\lfloor nx \rfloor, \lfloor ny \rfloor\} \in E_G; \\ 0 & \text{otherwise.} \end{cases}$$



A short course on graphons (1/2)

Graph function: with a graph G on vertex-set $\{0,\dots,n-1\}$, we associate its *rescaled adjacency matrix/pixel picture* $W_G : [0,1]^2 \to [0,1]$

$$W_G(x,y) = \begin{cases} 1 & \text{if } \{\lfloor nx \rfloor, \lfloor ny \rfloor\} \in E_G; \\ 0 & \text{otherwise.} \end{cases}$$

Idea: a sequence of graphs G_n if the associated functions W_{G_n} converge. . .

A short course on graphons (1/2)

Graph function: with a graph G on vertex-set $\{0,\dots,n-1\}$, we associate its *rescaled adjacency matrix/pixel picture* $W_G : [0,1]^2 \to [0,1]$

$$W_G(x,y) = \begin{cases} 1 & \text{if } \{\lfloor nx \rfloor, \lfloor ny \rfloor\} \in E_G; \\ 0 & \text{otherwise.} \end{cases}$$

Idea: a sequence of graphs G_n if the associated functions W_{G_n} converge... Yes, but

- it's better to use an adhoc norm

$$\| W \|_{\square} = \sup_{S,T \subseteq [0,1]} \left| \int_{S \times T} W(x,y) dxdy \right|.$$

- The function depends on the labelling of the vertices. We quotient by the relation

$$(W \sim W') \overset{\text{def}}{\iff} \exists \varphi \text{ Lebesgue-preserving} : W(\varphi(x), \varphi(y)) = W'(x,y).$$

A short course on graphons (2/2)

Theorem (Borgs, Chayes, Lovász, Sós, Vesztergombi, '08)

The following are equivalent

- W_{G_n} *converges to* W*;*
- *All subgraph proportions in* G_n *converge to some explicit functional of* W*, e.g.*

$$\frac{\#\ \text{edges in } G_n}{\binom{|V(G_n)|}{2}} \to \int_{[0,1]^2} W(x,y)dxdy,$$

$$\frac{\#\text{triangles in } G_n}{\binom{|V(G_n)|}{3}} \to \int_{[0,1]^2} W(x,y)W(y,z)W(z,x)dxdydz,$$

$$\dots$$

A short course on graphons (2/2)

Theorem (Borgs, Chayes, Lovász, Sós, Vesztergombi, '08)

The following are equivalent

- W_{G_n} *converges to* W;
- *All subgraph proportions in* G_n *converge to some explicit functional of* W, *e.g.*

$$\frac{\#\ \text{edges in } G_n}{\binom{|V(G_n)|}{2}} \to \int_{[0,1]^2} W(x,y)dxdy,$$

$$\frac{\#\text{triangles in } G_n}{\binom{|V(G_n)|}{3}} \to \int_{[0,1]^2} W(x,y)W(y,z)W(z,x)dxdydz,$$

$$\dots$$

- It gives a nice notion of limits for large graph;
- Introduced in 2006, and used a lot in large network analysis since then (even though real-life networks tend to be sparse).

First part

The static model

uniform random cographs

Limit of uniform random cographs

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

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

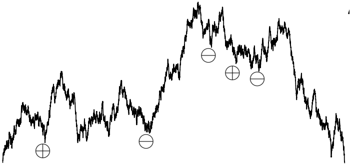
Limit of uniform random cographs

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

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

Construction of $\boldsymbol{W}_{Br}$: start from a Brownian excursion $\mathfrak{e}$ with i.i.d. balanced signs $(S(m))$ on local minima m of $\mathfrak{e}$ and set

$$\boldsymbol{W}_{Br}(x,y) = \begin{cases} 1 & \text{if } S\left(\operatorname{argmin}_{[x,y]} \mathfrak{e}\right) = \oplus; \\ 0 & \text{if } S\left(\operatorname{argmin}_{[x,y]} \mathfrak{e}\right) = \ominus. \end{cases}$$



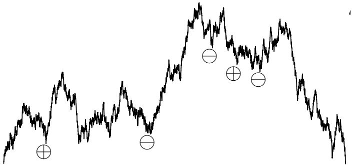
Limit of uniform random cographs

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

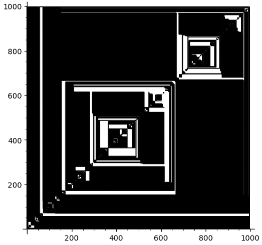
Let $\boldsymbol{G}_n$ be a uniform random (either labeled or unlabeled) cograph with n vertices. Then $W_{\boldsymbol{G}_n}$ converges in distribution to a random graphon $\boldsymbol{W}_{Br}$, which we call Brownian cographon.

Construction of $\boldsymbol{W}_{Br}$: start from a Brownian excursion $\mathfrak{e}$ with i.i.d. balanced signs ($S(m)$) on local minima m of $\mathfrak{e}$ and set

$$\boldsymbol{W}_{Br}(x,y) = \begin{cases} 1 & \text{if } S\left(\operatorname{argmin}_{[x,y]} \mathfrak{e}\right) = \oplus; \\ 0 & \text{if } S\left(\operatorname{argmin}_{[x,y]} \mathfrak{e}\right) = \ominus. \end{cases}$$

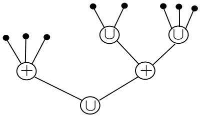


Adjacency matrix of a large uniform cograph (with a well-chosen order of vertices)



Heuristic for the theorem

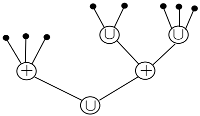
- A cograph G is encoded by a decorated tree T;



- Vertices in G correspond to leaves in T;
- Vertices v_1 and v_2 are connected if their youngest common ancestor is decorated by $\oplus$;

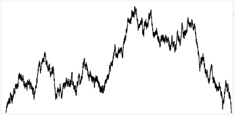
Heuristic for the theorem

- A cograph G is encoded by a decorated tree T;



- Vertices in G correspond to leaves in T;
- Vertices v_1 and v_2 are connected if their youngest common ancestor is decorated by $\oplus$;

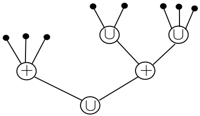
- The limit of T is Aldous' Continuum Random Tree T_∞, coded by a Brownian excursion $\mathfrak{e}$;



- Leaves of T_∞ form a measure 1 subset of $[0,1]$;
- Youngest common ancestor between x and y correspond to $\text{argmin}_{[x,y]}\, \mathfrak{e}$. Thus x, y are linked in $\boldsymbol{W}_{Br}$ if $S(\text{argmin}_{[x,y]}\, \mathfrak{e}) = \oplus$.

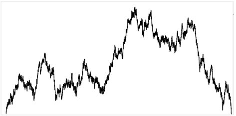
Heuristic for the theorem

- A cograph G is encoded by a decorated tree T;



- Vertices in G correspond to leaves in T;
- Vertices v_1 and v_2 are connected if their youngest common ancestor is decorated by $\oplus$;

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



- Leaves of T_∞ form a measure 1 subset of $[0,1]$;
- Youngest common ancestor between x and y correspond to $\mathrm{argmin}_{[x,y]}\,\mathfrak{e}$. Thus x, y are linked in $\boldsymbol{W}_{Br}$ if $S(\mathrm{argmin}_{[x,y]}\,\mathfrak{e}) = \oplus$.

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

Some consequences

Corollary

The edge proportion $|E(G_n)|/\binom{n}{2}$ in a uniform random cograph G_n with n vertices converges to a non-trivial random variable Λ.

Some consequences

Corollary

The edge proportion $|E(G_n)|/\binom{n}{2}$ in a uniform random cograph G_n with n vertices converges to a non-trivial random variable Λ.

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

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

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

The largest independent set in a uniform random cograph of size n has size $o_P(n)$.

Some consequences

Corollary

The edge proportion $|E(G_n)|/\binom{n}{2}$ *in a uniform random cograph* G_n *with* n *vertices converges to a non-trivial random variable* Λ.

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

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* $[0,1]$.

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

The largest independent set in a uniform random cograph of size n *has size* $o_P(n)$.

All these statements use an extension of the considered notion to graphons, some continuity property and some analysis of the Brownian cographon.

Erdős-Hajnal conjecture and the probabilistic version (1/2)

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
Without "H-free" constraints, optimal bound is $\log(n)$ (classical Ramsey theory).

Erdős-Hajnal conjecture and the probabilistic version (1/2)

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 and the probabilistic version (1/2)

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 a large family of graphs 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 and the probabilistic version (2/2)

Question (KMRS, '14)

Does there exists $b > 0$ such that a uniform random H-free graph G_n contains a homogeneous set of size bn with high probability?

Erdős-Hajnal conjecture and the probabilistic version (2/2)

Question (KMRS, '14)

Does there exists $b > 0$ such that a uniform random H-free graph G_n contains a homogeneous set of size bn with high probability?

Answer

No! We have seen that the largest independent set in G_n has size $o_P(n)$. By symmetry, the largest clique has also size $o_P(n)$, and, therefore, the largest independent set.

Second model

A dynamic model

with increasing size:

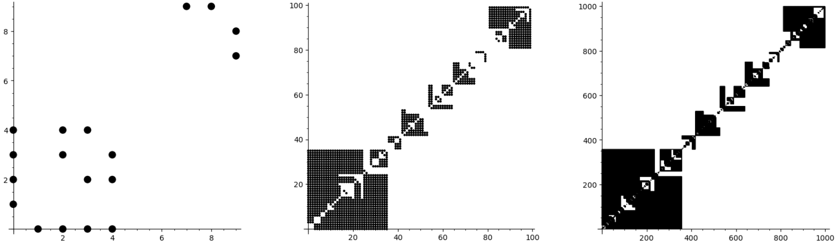
recursive random cographs

The model (reminder): H_n^p is constructed recursively by duplicating a uniform random vertex in H_{n-1}^p and connecting the two new vertices with probability p in $[0,1]$ (starting with $H_1^p = \bullet$).

A convergence result

Theorem (F., Rivera–Lopez, '23)

H_n^p converges almost surely *to a random graphon $\boldsymbol{W}_{rec}^p$, which we call the recursive cographon of paramter p.*



Adjacency matrices of $H_{10}^{1/2}$, $H_{100}^{1/2}$ and $H_{1000}^{1/2}$
in a single realization of the process $(H_n^{1/2})_{n\geq 1}$.

Construction of the limit

Let $(U_i)_{i \geq 0}$ be a sequence of i.i.d. uniform random variables in $[0,1]$, and $(S_i)_{i \geq 0}$ be a sequence of i.i.d. random signs in $\{\oplus, \ominus\}$.

For $x < y$ in $[0,1]$ let $i_{x,y}$ be the smallest index i s.t. $U_i \in [x,y)$. Then set

$$\boldsymbol{W}^p_{rec} = \begin{cases} 1 & \text{if } S_{i_{x,y}} = \oplus; \\ 0 & \text{otherwise.} \end{cases}$$

Construction of the limit

Let $(U_i)_{i\geq 0}$ be a sequence of i.i.d. uniform random variables in $[0,1]$, and $(S_i)_{i\geq 0}$ be a sequence of i.i.d. random signs in $\{\oplus,\ominus\}$.

For $x<y$ in $[0,1]$ let $i_{x,y}$ be the smallest index i s.t. $U_i \in [x,y)$. Then set

$$\boldsymbol{W}^p_{rec} = \begin{cases} 1 & \text{if } S_{i_{x,y}} = \oplus; \\ 0 & \text{otherwise.} \end{cases}$$

Comparison with $\boldsymbol{W}_{Br}$: no Brownian structure behind, the minima of the Brownian excursion are replaced by the U_i.

Note. Such differences between uniform/recursive structures had been observed earlier on other objects: trees, triangulations of the disk, ...

Uniform vs Recursive

Proposition (F., Rivera–Lopez, '23)

The distributions of the Brownian and recursive cographon are mutually singular.

i.e. we can exhibit a set X of graphons such that

$$\mathbb{P}[W_{Br} \in X] = 0 = 1 - \mathbb{P}[W_{rec}^{p} \in X].$$

Uniform vs Recursive

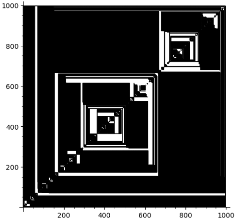
Proposition (F., Rivera–Lopez, '23)

The distributions of the Brownian and recursive cographon are mutually singular.

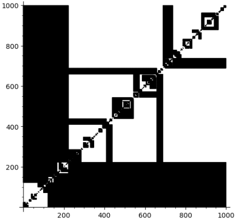
i.e. we can exhibit a set X of graphons such that

$$\mathbb{P}[W_{Br} \in X] = 0 = 1 - \mathbb{P}[W_{rec}^{p} \in X].$$

In fact $X = \left\{ W \text{ s.t. } \exists \varepsilon : \text{W constant on } [0,\varepsilon] \times [1-\varepsilon, 1] \right\}$ works!



uniform



recursive

Proof strategy for $H_n^p \xrightarrow{a.s.} W_{rec}^p$.

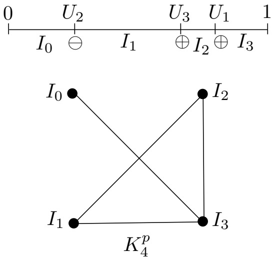
Difficulty: since we want a.s. convergence, we need to realize the process $(H_n^p)_{n\geq 1}$ et the limit W_{rec}^p on the same probability space.

Proof strategy for $H_n^p \xrightarrow{a.s.} W_{rec}^p$.

Difficulty: since we want a.s. convergence, we need to realize the process $(H_n^p)_{n\geq 1}$ et the limit W_{rec}^p on the same probability space.

Solution: take U_i, S_i as above, and for each n, let K_n^p be the following graph

- vertices $\leftrightarrow$ intervals defined by cutting $[0,1]$ at $U_1, \dots, U_{n-1}$;
- edges: $\{I, J\}$ is an edge iff the U with smallest index between I and J has a $\oplus$ sign.

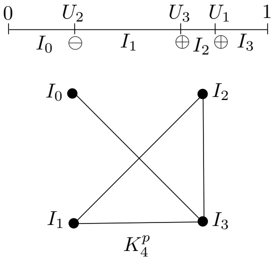


Proof strategy for $H_n^p \xrightarrow{a.s.} W_{rec}^p$.

Difficulty: since we want a.s. convergence, we need to realize the process $(H_n^p)_{n\geq 1}$ et the limit W_{rec}^p on the same probability space.

Solution: take U_i, S_i as above, and for each n, let K_n^p be the following graph

- vertices $\leftrightarrow$ intervals defined by cutting $[0,1]$ at $U_1, \dots, U_{n-1}$;
- edges: $\{I, J\}$ is an edge iff the U with smallest index between I and J has a $\oplus$ sign.



Lemma 1

$$K_n^p \xrightarrow{a.s.} W_{rec}^p.$$

Lemma 2

$$(K_n^p)_{n\geq 1} \stackrel{law}{=} (H_n^p)_{n\geq 1}.$$

Third model

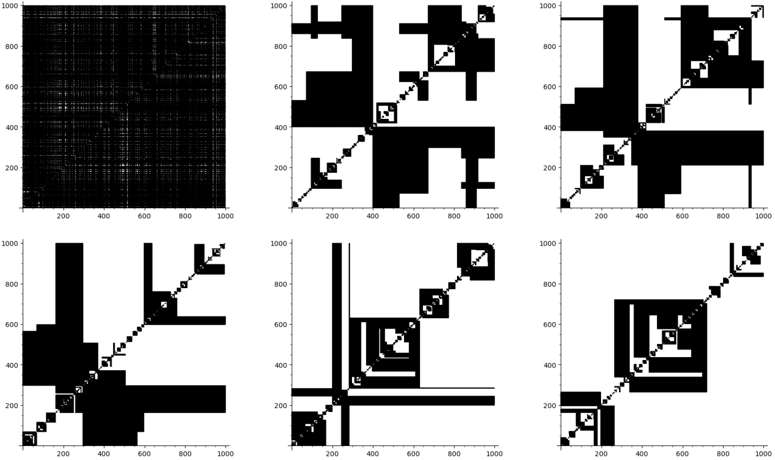
A dynamic model

with constant size:

up down chain on cographs

The model (reminder) : $X_n^p(t)$ obtained by duplicating a uniform random vertex (connecting the two new vertices with proba p) and then deleting a uniform random vertex in $X_n^p(t-1)$,
starting with *any* graph $X_n^p(0)$ with n vertices.

A simulation of the updown chain



Simulation of the updown chain for $n = 1000$, $p = 1/2$, and $t \in \{0, 1, 2, 3, 4, 5\} \cdot 50000$.

Asymptotics of up-down chains

Theorem (F., Rivera-Lopez, in preparation)

Fix a sequence $X_n^p(0)$ *of (possibly random) initial configurations, and assume that* $X_n^p(0)$ *converges (in distribution) to a (possibly random) graphon* $W(0)$.
Then $X_n^p(\lfloor n^2 t \rfloor)$ *converges in distribution to some non-trivial Feller process* F^p *with initial distribution* $W(0)$ *in the Skorohod space* $\mathscr{D}([0,+\infty)$.

Asymptotics of up-down chains

Theorem (F., Rivera-Lopez, in preparation)

Fix a sequence $X_n^p(0)$ of (possibly random) initial configurations, and assume that $X_n^p(0)$ converges (in distribution) to a (possibly random) graphon $W(0)$.
Then $X_n^p(\lfloor n^2 t \rfloor)$ converges in distribution to some non-trivial Feller process F^p with initial distribution $W(0)$ in the Skorohod space $\mathscr{D}([0,+\infty)$.

What we know about F^p?

- Subgraph density of non-cograph patterns decay exponentially fast with an explicit rate;
- The stationary distribution of F^p is the random recursive cograph W_{rec}^p;
- Hopefully more soon (ergodicity, path continuity, ...)...