Local analysis of matchings on sparse graphs

Asymptotic local law of matchings on random graphs

Mike Liu[1] Nathanaël Enriquez[2] Laurent Ménard[3]
Vianney Perchet[4]

[1]CREST
ENSAE Paris

[2]LMO
Université Paris-Saclay

[3]Modal'X
Université Paris-Nanterre

[4]CREST
ENSAE Paris

December 2023

Table of Contents

Table of Contents

Erdös-Renyi model

Let $n \in \mathbb{N}^*$, $p \in [0,1]$

Definition

$G_{n,p}$ is the random graph on $V = \{1, 2, ..., n\}$ such that independently on every (i,j), $\mathbb{P}((i,j) \in E) = p$.

Erdös-Renyi model

Let $n \in \mathbb{N}^*$, $p \in [0, 1]$

Definition

$G_{n,p}$ is the random graph on $V = \{1, 2, ..., n\}$ such that independently on every (i, j), $\mathbb{P}((i, j) \in E) = p$.

Let $\mathcal{W}$ be some law on $\mathbb{R}_+$ with finite expectation. (for example, uniform on $[0, 1]$).

Definition

$G_{n,p,\mathcal{W}}$ is the weighted random graph generated on $G_{n,p} = (V, E)$ by drawing $(W(e))_{e \in E}$ i.i.d of law $\mathcal{W}$.

Maximum matching, Minimum cost maximum matching

Let $G = (V, E, w)$ a weighted graph.

Definition

A matching on G is a subgraph $\mathcal{M}$ of G such that $deg(v) \leq 1$ for all $v \in V$

Maximum matching, Minimum cost maximum matching

Let $G = (V, E, w)$ a weighted graph.

Definition

A matching on G is a subgraph $\mathcal{M}$ of G such that $deg(v) \leq 1$ for all $v \in V$

- We say that it is of maximum weight if the sum of the weight on its edges is maximal among all matchings.

Maximum matching, Minimum cost maximum matching

Let $G = (V, E, w)$ a weighted graph.

Definition

A matching on G is a subgraph $\mathcal{M}$ of G such that $deg(v) \leq 1$ for all $v \in V$

- We say that it is of maximum weight if the sum of the weight on its edges is maximal among all matchings.
- We say that a matching is of maximum cardinality of its size is maximal among all matchings.

Maximum matching, Minimum cost maximum matching

Let $G = (V, E, w)$ a weighted graph.

Definition

A matching on G is a subgraph $\mathcal{M}$ of G such that $deg(v) \leq 1$ for all $v \in V$

- We say that it is of maximum weight if the sum of the weight on its edges is maximal among all matchings.
- We say that a matching is of maximum cardinality of its size is maximal among all matchings.
- We say that it is a minimum cost maximum matching if the sum of the weight on its edge is minimum among matchings of maximum cardinality.

Local convergence

Local convergence

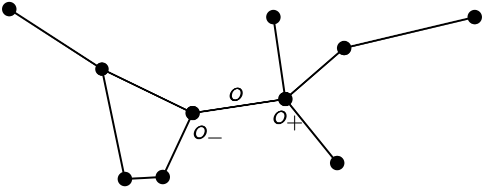
By distinguishing an edge in G, we define a rooted graph $G' = (V, E, w, o)$.

Local convergence

By distinguishing an edge in G, we define a rooted graph $G' = (V, E, w, o)$.
Let $H \in \mathbb{N}$, we define $N_H(G') (= N_H(o))$ the $H-$neighboorhood of the root $o = (o_-, o_+)$ in G' by the set of edges at distance less than H from the root and the set of vertices at distance less than $H - 1$ from o_- or o_+.

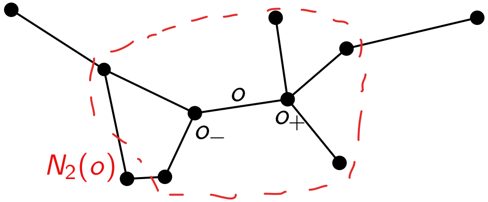
Local convergence

By distinguishing an edge in G, we define a rooted graph $G' = (V, E, w, o)$.
Let $H \in \mathbb{N}$, we define $N_H(G')(= N_H(o))$ the $H-$neighboorhood of the root $o = (o_-, o_+)$ in G' by the set of edges at distance less than H from the root and the set of vertices at distance less than $H - 1$ from o_- or o_+.



Local convergence

By distinguishing an edge in G, we define a rooted graph $G' = (V, E, w, o)$.
Let $H \in \mathbb{N}$, we define $N_H(G')(= N_H(o))$ the $H-$neighboorhood of the root $o = (o_-, o_+)$ in G' by the set of edges at distance less than H from the root and the set of vertices at distance less than $H - 1$ from o_- or o_+.



Local convergence (2)

Local convergence (2)

Let $(G_n)_{n\in\mathbb{N}}$ be a sequence of rooted random graphs.

Definition

We say that G_n converges locally to G if for every $H \in \mathbb{N}$, the $H-$neighborhood of the root of G_n converges in law to the $H-$neighborhood of the root of G.

Table of Contents

Context for main theorems

Context for main theorems

Assume that $\mathcal{W}$ is a continuous law with no atoms.
W.h.p, the maximum weight matching in $G_{n,p,\mathcal{W}}$ (resp. the minimum cost maximum matching) is uniquely defined, we call them $\mathcal{M}_n^{opt}$ (resp. $\mathcal{M}_n^{mc\,opt}$).

Context for main theorems

Assume that $\mathcal{W}$ is a continuous law with no atoms.
W.h.p, the maximum weight matching in $G_{n,p,\mathcal{W}}$ (resp. the minimum cost maximum matching) is uniquely defined, we call them $\mathcal{M}_n^{opt}$ (resp. $\mathcal{M}_n^{mc\,opt}$).
Fix $c > 0$, we are interested in the sparse regime where $p = \frac{c}{n}$.

Context for main theorems

Assume that $\mathcal{W}$ is a continuous law with no atoms.
W.h.p, the maximum weight matching in $G_{n,p,\mathcal{W}}$ (resp. the minimum cost maximum matching) is uniquely defined, we call them $\mathcal{M}_n^{opt}$ (resp. $\mathcal{M}_n^{mc\,opt}$).
Fix $c > 0$, we are interested in the sparse regime where $p = \frac{c}{n}$.
For any n, let us root $G_{n,\frac{c}{n},\mathcal{W}}$ by picking o_n uniformly among the edges.

Context for main theorems

Assume that $\mathcal{W}$ is a continuous law with no atoms.
W.h.p, the maximum weight matching in $G_{n,p,\mathcal{W}}$ (resp. the minimum cost maximum matching) is uniquely defined, we call them $\mathcal{M}_n^{opt}$ (resp. $\mathcal{M}_n^{mc\,opt}$).
Fix $c > 0$, we are interested in the sparse regime where $p = \frac{c}{n}$.
For any n, let us root $G_{n,\frac{c}{n},\mathcal{W}}$ by picking o_n uniformly among the edges.
Let us call $\mathcal{T}_{\mathcal{W}}$ the weighted edge-rooted Galton-Watson tree of reproduction law $\mathcal{P}(c)$ with i.i.d weights on edges drawn with law $\mathcal{W}$.
It is known that $G_{n,\frac{c}{n},\mathcal{W}}$ converge locally to $\mathcal{T}_{\mathcal{W}}$.

Main theorems

Main theorems

Theorem (2024, Liu et al.)

$((G_{(n,\frac{c}{n},\mathcal{W})}, o_n), (\mathcal{M}_n^{opt}, o_n))$ converge locally to $(\mathcal{T}_{\mathcal{W}}, \mathcal{M})$ where $\mathcal{M}$ is a matching on $\mathcal{T}_{\mathcal{W}}$ that can be described.

Main theorems

Theorem (2024, Liu et al.)

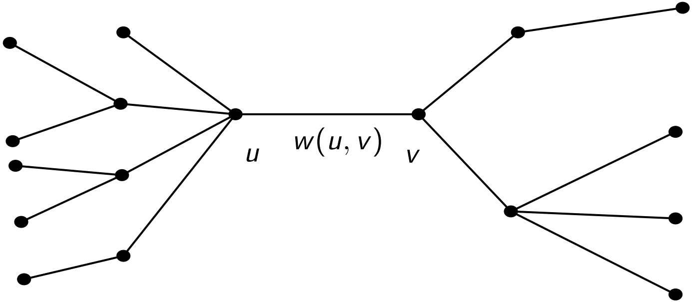
$((G_{(n,\frac{c}{n},\mathcal{W})}, o_n), (\mathcal{M}_n^{opt}, o_n))$ converge locally to $(\mathcal{T}_{\mathcal{W}}, \mathcal{M})$ where $\mathcal{M}$ is a matching on $\mathcal{T}_{\mathcal{W}}$ that can be described.

Theorem 2 (2024 Liu et al.)

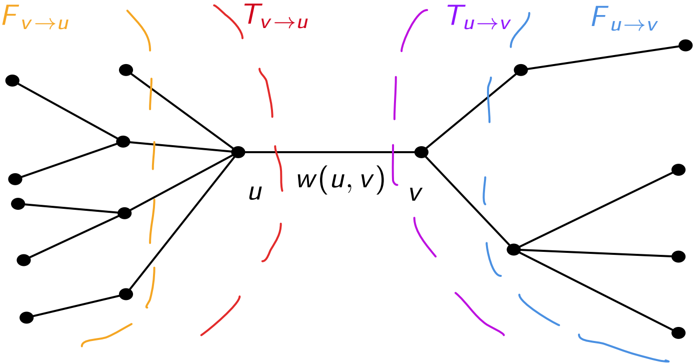
$((G_{(n,\frac{c}{n},\mathcal{W})}, o_n), (\mathcal{M}_n^{opt\ mc}, o_n))$ converge locally to $(\mathcal{T}_{\mathcal{W}}, \mathcal{M}^{mc})$ where $\mathcal{M}^{mc}$ is a matching on $\mathcal{T}_{\mathcal{W}}$ that can be described.

Table of Contents

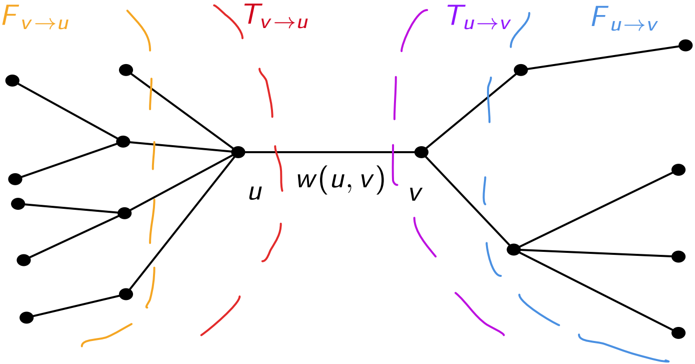
Dynamic programming heuristic



Dynamic programming heuristic



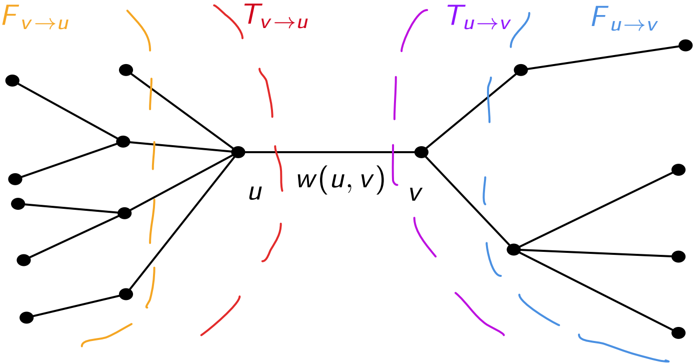
Dynamic programming heuristic



$Z(v, u) :=$
$OPT(T_{v \to u}) - OPT(F_{v \to u})$.

$Z(u, v) :=$
$OPT(T_{u \to v}) - OPT(F_{u \to v})$.

Dynamic programming heuristic

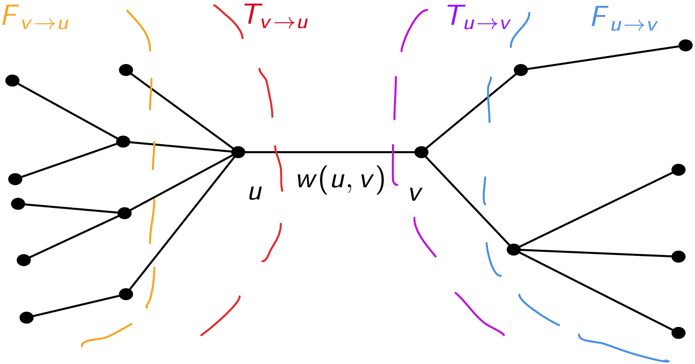


$Z(v,u) :=$
$OPT(T_{v\to u}) - OPT(F_{v\to u})$.

$Z(u,v) :=$
$OPT(T_{u\to v}) - OPT(F_{u\to v})$.

$OPT_{\text{forcing }(u,v)}(T) - OPT_{\text{without }(u,v)}(T)$

Dynamic programming heuristic

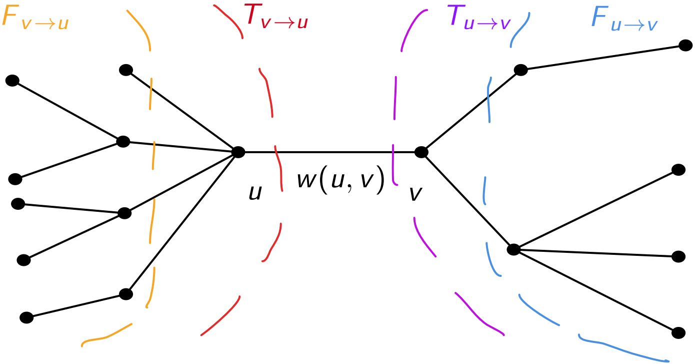


$Z(v,u) :=$
$OPT(T_{v\to u}) - OPT(F_{v\to u})$.

$Z(u,v) :=$
$OPT(T_{u\to v}) - OPT(F_{u\to v})$.

$$OPT_{\text{forcing }(u,v)}(T) - OPT_{\text{without }(u,v)}(T) = OPT(F_{u\to v}) + OPT(F_{v\to u}) + w(u,v) - OPT(T_{v\to u}) - OPT(T_{u\to v})$$

Dynamic programming heuristic

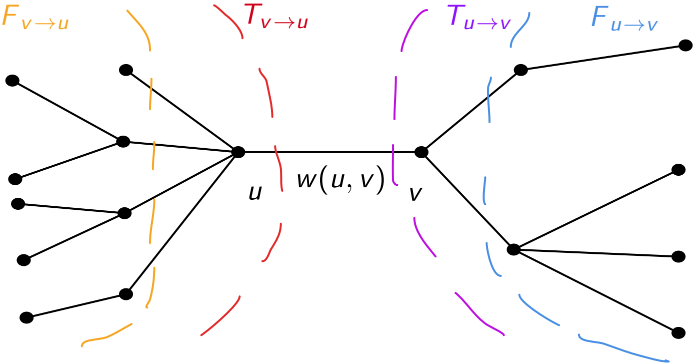


$Z(v,u) :=$
$OPT(T_{v\to u}) - OPT(F_{v\to u})$.

$Z(u,v) :=$
$OPT(T_{u\to v}) - OPT(F_{u\to v})$.

$OPT_{\text{forcing }(u,v)}(T) - OPT_{\text{without }(u,v)}(T) = OPT(F_{u\to v}) + OPT(F_{v\to u}) + w(u,v)$
$-OPT(T_{v\to u}) - OPT(T_{u\to v}) = w(u,v) - Z(u,v) - Z(v,u)$.

Dynamic programming heuristic



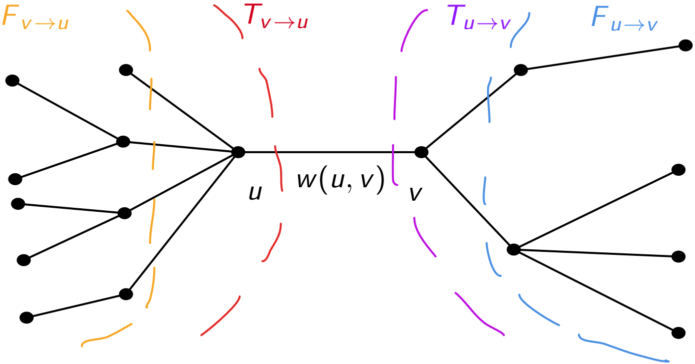
$Z(v,u) :=$ $OPT(T_{v\to u}) - OPT(F_{v\to u})$.

$Z(u,v) :=$ $OPT(T_{u\to v}) - OPT(F_{u\to v})$.

$$OPT_{\text{forcing }(u,v)}(T) - OPT_{\text{without }(u,v)}(T) = OPT(F_{u\to v}) + OPT(F_{v\to u}) + w(u,v)$$
$$- OPT(T_{v\to u}) - OPT(T_{u\to v}) = w(u,v) - Z(u,v) - Z(v,u).$$

$(u,v) \in OPT$

Dynamic programming heuristic



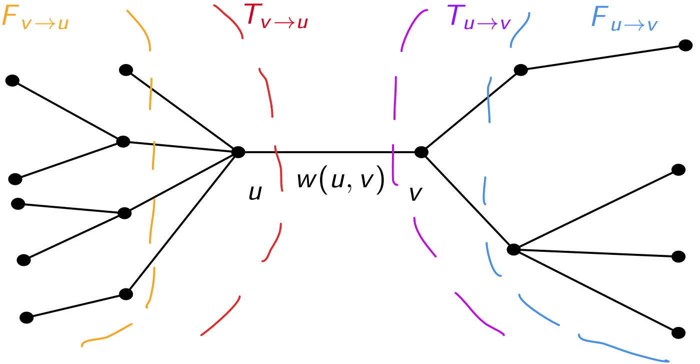
$Z(v,u) :=$
$OPT(T_{v\to u}) - OPT(F_{v\to u})$.

$Z(u,v) :=$
$OPT(T_{u\to v}) - OPT(F_{u\to v})$.

$$OPT_{\text{forcing }(u,v)}(T) - OPT_{\text{without }(u,v)}(T) = OPT(F_{u\to v}) + OPT(F_{v\to u}) + w(u,v) - OPT(T_{v\to u}) - OPT(T_{u\to v}) = w(u,v) - Z(u,v) - Z(v,u).$$

$$(u,v) \in OPT \quad \Leftrightarrow OPT_{\text{forcing }(u,v)}(T) - OPT_{\text{without }(u,v)}(T) > 0$$

Dynamic programming heuristic



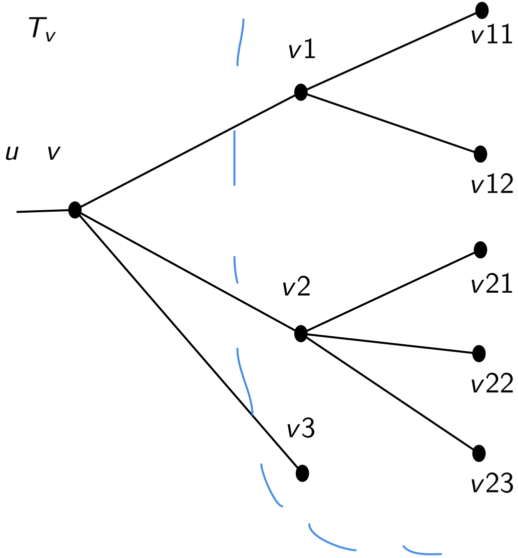
$Z(v,u) :=$
$OPT(T_{v \to u}) - OPT(F_{v \to u})$.

$Z(u,v) :=$
$OPT(T_{u \to v}) - OPT(F_{u \to v})$.

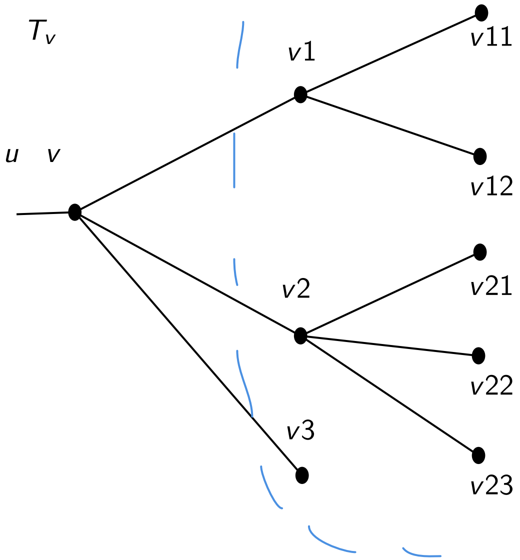
$$OPT_{\text{forcing }(u,v)}(T) - OPT_{\text{without }(u,v)}(T) = OPT(F_{u \to v}) + OPT(F_{v \to u}) + w(u,v)$$
$$-OPT(T_{v \to u}) - OPT(T_{u \to v}) = w(u,v) - Z(u,v) - Z(v,u).$$

$$(u,v) \in OPT \Leftrightarrow OPT_{\text{forcing }(u,v)}(T) - OPT_{\text{without }(u,v)}(T) > 0$$
$$\Leftrightarrow Z(u,v) + Z(v,u) < w(u,v).$$

Recursive equation on heuristic

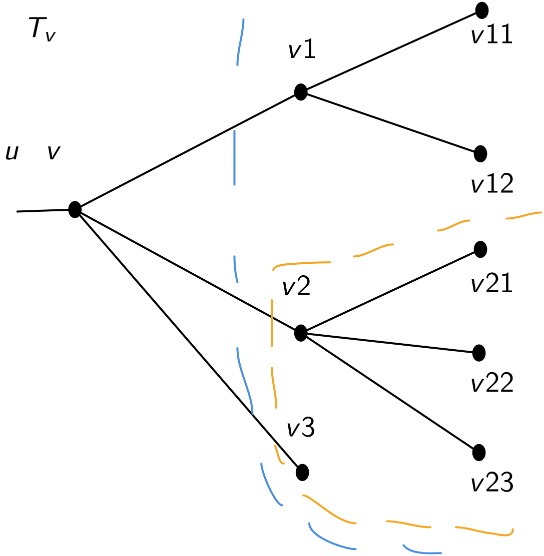


Recursive equation on heuristic



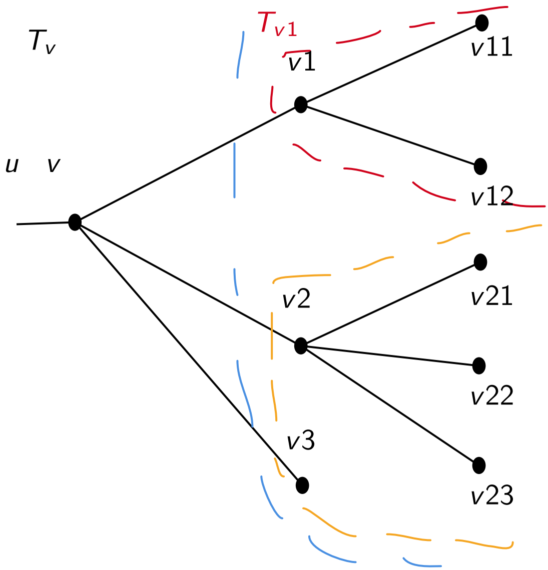
Assume that $OPT(T_v)$ picks $v1$,

Recursive equation on heuristic



Assume that $OPT(T_v)$ picks v1, OPT(orange) is in common between $OPT(T_v)$ and $OPT(F_{u \to v})$.

Recursive equation on heuristic

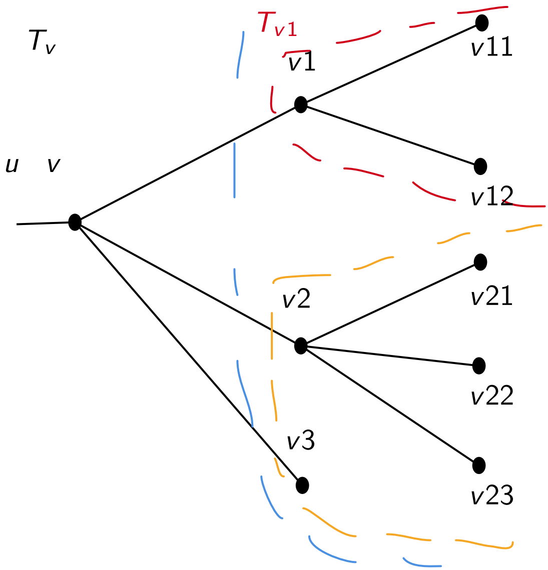


Assume that $OPT(T_v)$ picks v1, OPT(orange) is in common between $OPT(T_v)$ and $OPT(F_{u \to v})$.

In red, $OPT(T_v)$ optimizes $OPT(T_{v11})$ and $OPT(T_{v12})$.

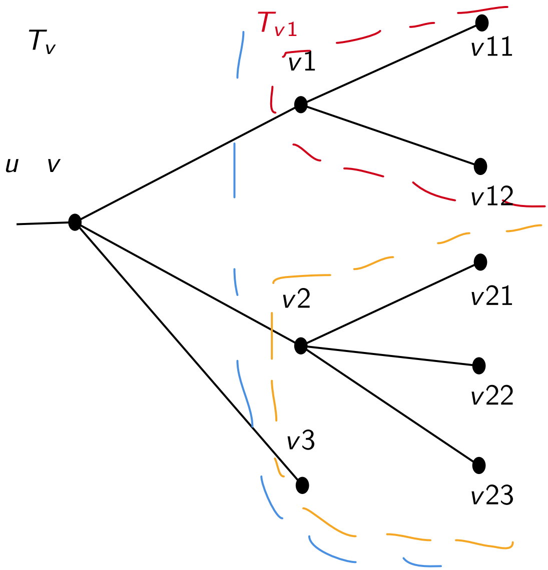
$OPT(F_{u \to v})$ optimizes $OPT(T_{v1})$.

Recursive equation on heuristic



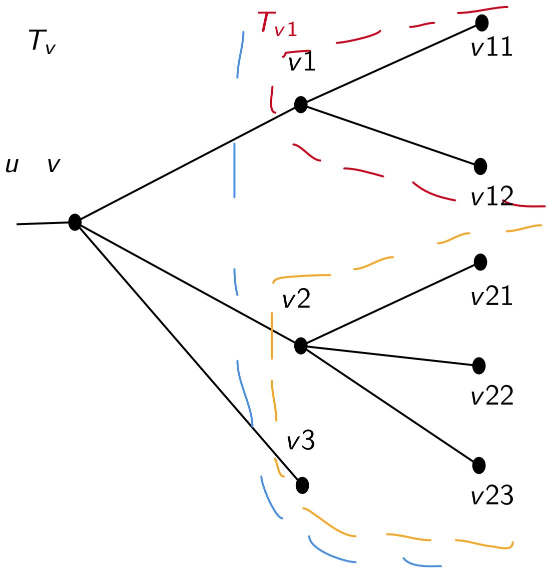
Assume that $OPT(T_v)$ *picks* $v1$,
OPT(*orange*) *is in common between* $OPT(T_v)$ *and* $OPT(F_{u\to v})$.
In red, $OPT(T_v)$ *optimizes* $OPT(T_{v11})$ *and* $OPT(T_{v12})$.
$OPT(F_{u\to v})$ *optimizes* $OPT(T_{v1})$.
Hence $Z(u,v) = OPT(T_v) - OPT(F_{u\to v})$
$= w(v, v1) - OPT(T_{v1}) + OPT(T_{v11})$
$+ OPT(T_{v12})$

Recursive equation on heuristic



Assume that $OPT(T_v)$ *picks* $v1$,
OPT(orange) *is in common between*
$OPT(T_v)$ *and* $OPT(F_{u\to v})$.
In red, $OPT(T_v)$ *optimizes* $OPT(T_{v11})$
and $OPT(T_{v12})$.
$OPT(F_{u\to v})$ *optimizes* $OPT(T_{v1})$.
Hence $Z(u,v) = OPT(T_v) - OPT(F_{u\to v})$
$= \ w(v, v1) - OPT(T_{v1}) + OPT(T_{v11})$
$+ OPT(T_{v12})$
$= w(v, v1) - Z(v, v1)$

Recursive equation on heuristic



Assume that $OPT(T_v)$ *picks* $v1$,
$OPT($*orange*$)$ *is in common between*
$OPT(T_v)$ *and* $OPT(F_{u \to v})$.
In red, $OPT(T_v)$ *optimizes* $OPT(T_{v11})$
and $OPT(T_{v12})$.
$OPT(F_{u \to v})$ *optimizes* $OPT(T_{v1})$.
Hence $Z(u, v) = OPT(T_v) - OPT(F_{u \to v})$
$= w(v, v1) - OPT(T_{v1}) + OPT(T_{v11})$
$+ OPT(T_{v12})$
$= w(v, v1) - Z(v, v1)$
$v1$ *has to be maximal for this quantity hence :*

$$Z(u, v) = max_{w \sim v, w \neq u}(0, w(v, w) - Z(v, w))$$

Equation in law

Equation in law

Assume there exists a law $\mathcal{L}$ such that for Z, Z_k i.i.d of law $\mathcal{L}$, w_k of law $\mathcal{W}$ and $K \sim \mathcal{P}(c)$ all mutually independent such that:

$$Z \stackrel{\mathcal{D}}{=} \max(0, \max_{1 \leq k \leq K} (w_k - Z_k)).$$

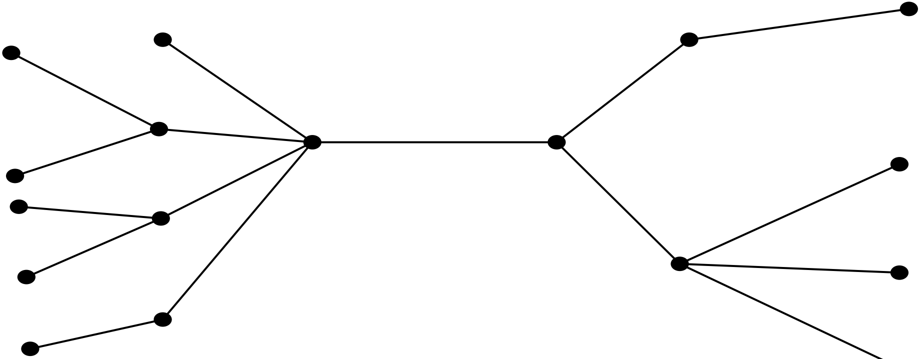
Equation in law

Assume there exists a law $\mathcal{L}$ such that for Z, Z_k i.i.d of law $\mathcal{L}$, w_k of law $\mathcal{W}$ and $K \sim \mathcal{P}(c)$ all mutually independent such that:

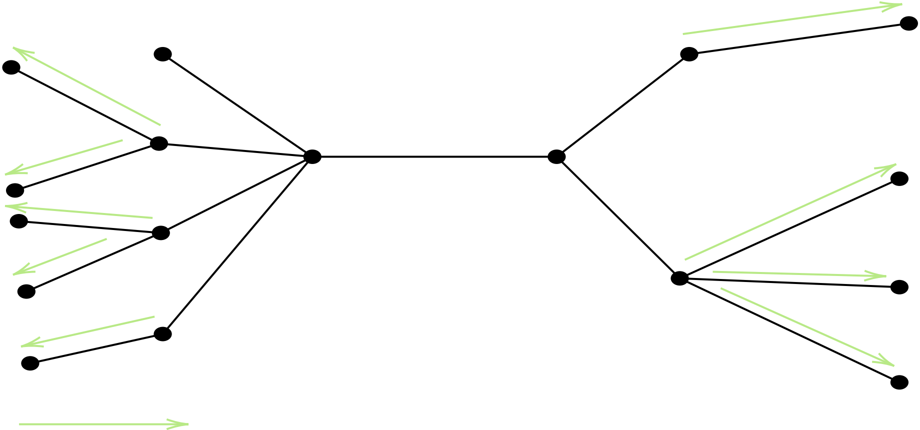
$$Z \overset{\mathcal{D}}{=} \max(0, \max_{1 \leq k \leq K} (w_k - Z_k)).$$

Fix $H \in \mathbb{N}$. We define Z in $N_H(o)$ of $T_{\mathcal{W}}$ in a way such that the law is compatible as H increases.

Existence of Z on T and Kolmogorov

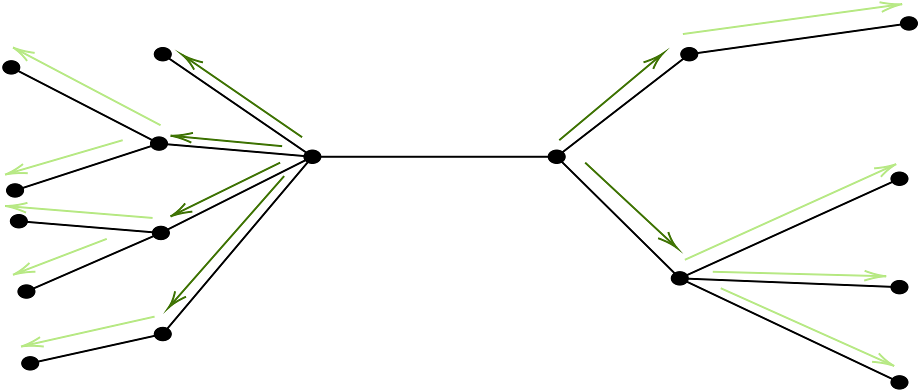


Existence of Z on T and Kolmogorov



Z *drawn i.i.d with law* $\mathcal{L}$

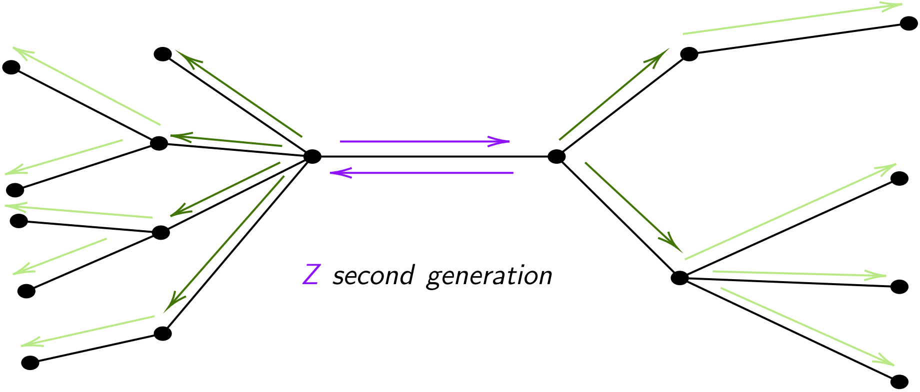
Existence of Z on T and Kolmogorov



Z *drawn i.i.d with law* $\mathcal{L}$

Z *first generation constructed by recursion*

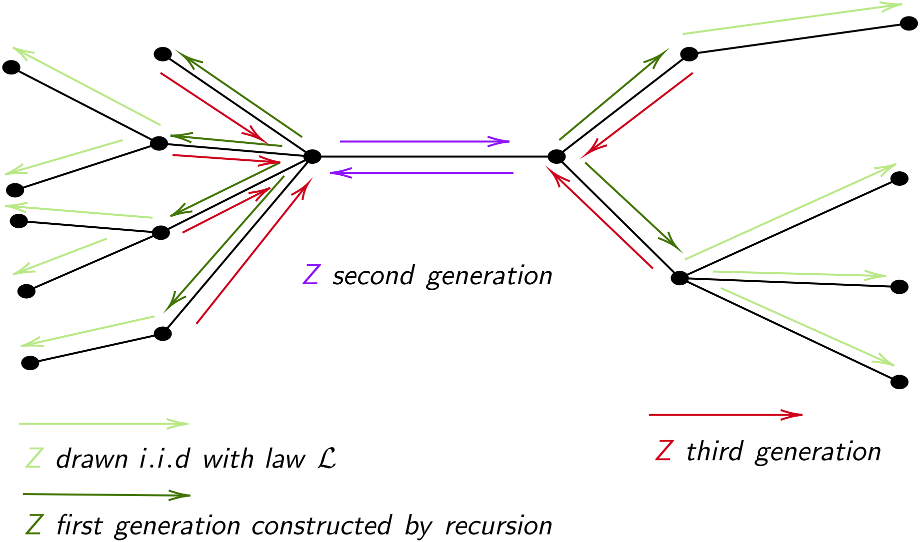
Existence of Z on T and Kolmogorov



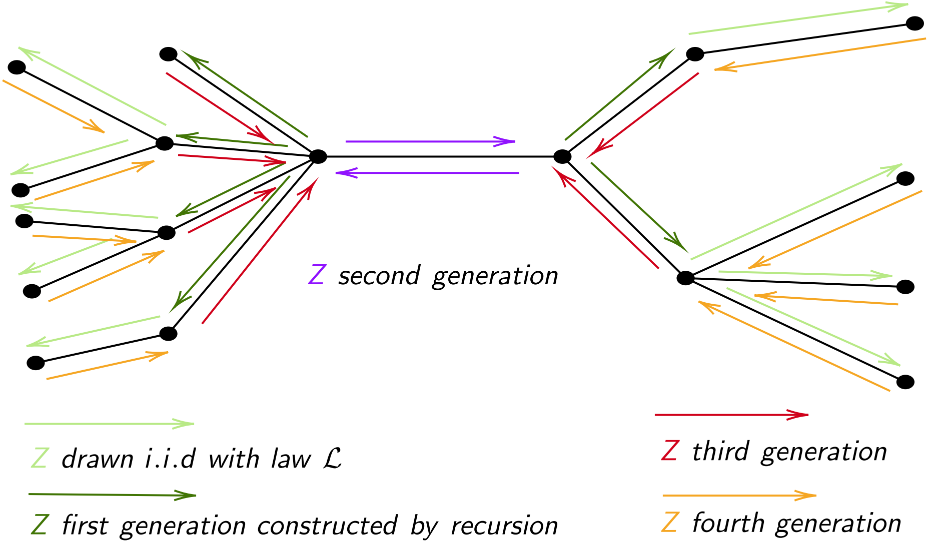
Z *drawn i.i.d with law* $\mathcal{L}$

Z *first generation constructed by recursion*

Existence of Z on T and Kolmogorov



Existence of Z on T and Kolmogorov



Description of $\mathcal{M}$

We now define $\mathcal{M}$ as the set of (u, v) such that

$$Z(u, v) + Z(v, u) < W(u, v).$$

This defines the limit law in Theorem 1.

Description of $\mathcal{M}^{mc}$

Description of $\mathcal{M}^{mc}$

Replace $\mathcal{W}$ by $\mathcal{W}_\epsilon$ the law of $1 - \epsilon W$ for $W \sim \mathcal{W}$ and $\epsilon > 0$.

Description of $\mathcal{M}^{mc}$

Replace $\mathcal{W}$ by $\mathcal{W}_\epsilon$ the law of $1 - \epsilon W$ for $W \sim \mathcal{W}$ and $\epsilon > 0$.
As ϵ goes to 0 and through renormalization methods, it is possible to show that we need to replace Z by:
$\Phi = (\Phi_0, \Phi_1)$ when $c \leq e$,
$\Phi = (\Phi_0, \Phi_{\frac{1}{2}}, \Phi_1)$ when $c > e$. with:

$$\Phi_0(u,v) = \max(0, \max_{\substack{w \sim v \\ w \neq u}}(-w(v,w) - \Phi_1(v,w))),$$

$$\Phi_1(u,v) = \max_{\substack{w \sim v \\ w \neq u}}(-w(v,w) - \Phi_0(v,w)),$$

$$\Phi_{\frac{1}{2}}(u,v) = \max_{\substack{w \sim v \\ w \neq u}}(-w(v,w) - \Phi_{\frac{1}{2}}(v,w)).$$

Description of $\mathcal{M}^{mc}(2)$

We then define $\mathcal{M}^{mc}$ as the set of (u, v) such that :

$$\Phi_0(u, v) + \Phi_1(v, u) < -w(u, v)$$

or

$$\Phi_1(u, v) + \Phi_0(v, u) < -w(u, v)$$

or

$$\Phi_{\frac{1}{2}}(u, v) + \Phi_{\frac{1}{2}}(v, u) < -w(u, v).$$

Table of Contents

Asymptotic costs

Asymptotic costs

Asymptotic cost of minimum cost maximum matching on Erdös-Renyi Graph

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\sum_{e\in\mathcal{M}_n^{\text{opt mc}}} w(e)\right] = \frac{c}{2}\mathbb{E}\left[w(o)\mathbb{1}_{o\in\mathcal{M}^{mc}}\right]$$

Asymptotic costs

Asymptotic cost of minimum cost maximum matching on Erdös-Renyi Graph

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\sum_{e\in\mathcal{M}_n^{\text{opt mc}}} w(e)\right] = \frac{c}{2}\mathbb{E}\left[w(o)\mathbb{1}_{o\in\mathcal{M}^{mc}}\right]$$

We can also recover the known Karp-Sipser formula of maximum matching size

Corollary

$$\lim_{n\to\infty} \frac{1}{n}\left(\textit{Size of maximum matching on } G_{n,\frac{c}{n}}\right)$$

$$= \lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\sum_{e\in\mathcal{M}_n^{opt\ mc}} 1\right] = \frac{c}{2}\mathbb{E}\left[\mathbb{1}_{o\in\mathcal{M}^{mc}}\right]$$

Conditioning on local geometry

Conditioning on local geometry

For either $\mathcal{M}_n = \mathcal{M}_n^{opt}$ or $\mathcal{M}_n^{\text{opt mc}}$, we can also compute any conditioning on local property such as:

$$\lim_{n\to\infty} \mathbb{P}\left(e \text{ is matched by } \mathcal{M}_n | w(e) \leq x\right)$$

for some fixed x,

Conditioning on local geometry

For either $\mathcal{M}_n = \mathcal{M}_n^{opt}$ or $\mathcal{M}_n^{\text{opt mc}}$, we can also compute any conditioning on local property such as:

$$\lim_{n\to\infty} \mathbb{P}\left(e \text{ is matched by } \mathcal{M}_n | w(e) \leq x\right)$$

for some fixed x, or

$$\lim_{n\to\infty} \mathbb{P}\left(e \text{ is matched by } \mathcal{M}_n | \text{the degrees of } e \text{ are 3 and 4}\right)$$

Conditioning on local geometry

For either $\mathcal{M}_n = \mathcal{M}_n^{opt}$ or $\mathcal{M}_n^{\text{opt mc}}$, we can also compute any conditioning on local property such as:

$$\lim_{n\to\infty} \mathbb{P}\left(e \text{ is matched by } \mathcal{M}_n | w(e) \leq x\right)$$

for some fixed x, or

$$\lim_{n\to\infty} \mathbb{P}\left(e \text{ is matched by } \mathcal{M}_n | \text{the degrees of } e \text{ are 3 and 4}\right)$$

In general, fix some $H \in \mathbb{N}$, if $\mathcal{G}$ is a collection of rooted weighted graphs and $\mathcal{G}_H$ the collection of their $H-$neighborhood of their roots, we can compute

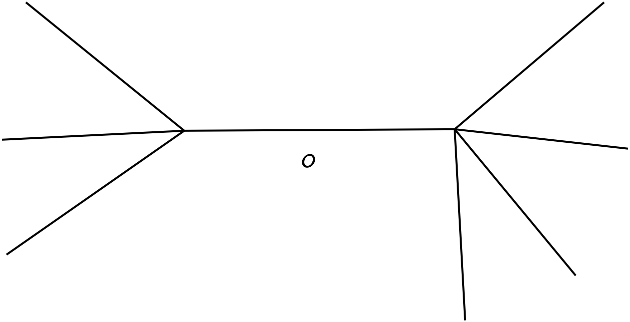
$$\lim_{n\to\infty} \mathbb{P}\left(e \text{ is matched by } \mathcal{M}_n | N_H(e_n) \in \mathcal{G}_H\right).$$

Conditioning on local geometry (2)

For the second limit, the limit will be the probability that $o \in \mathcal{M}$ with the following configuration (with independent weights drawn on edges) :

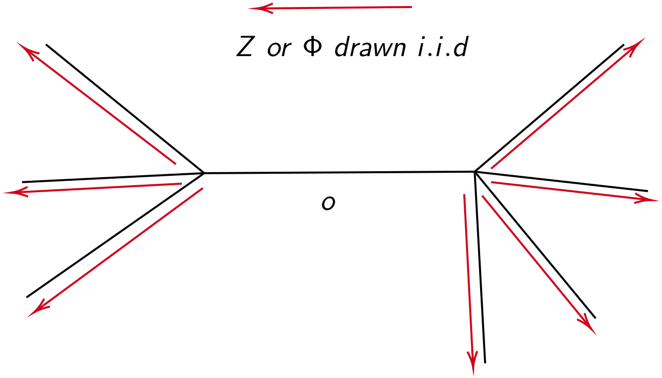
Conditioning on local geometry (2)

For the second limit, the limit will be the probability that $o \in \mathcal{M}$ with the following configuration (with independent weights drawn on edges) :



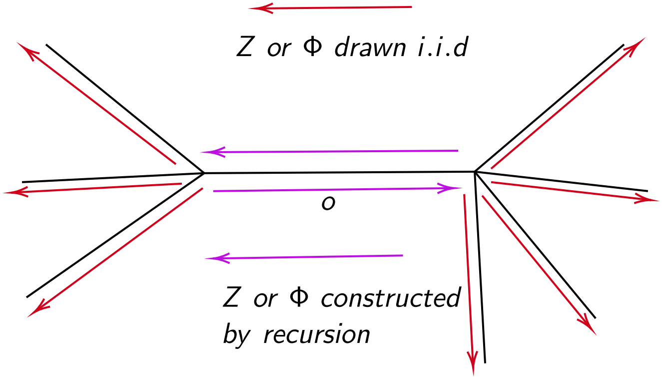
Conditioning on local geometry (2)

For the second limit, the limit will be the probability that $o \in \mathcal{M}$ with the following configuration (with independent weights drawn on edges) :



Conditioning on local geometry (2)

For the second limit, the limit will be the probability that $o \in \mathcal{M}$ with the following configuration (with independent weights drawn on edges) :



Surprising statistics

Surprisingly, there are some quantities that do not depend on the law $\mathcal{W}$. It is possible to compute

$$\lim_{n\to\infty} \mathbb{P}\left(e = (u,v) \text{ is matched by } \mathcal{M}_n^{\text{opt mc}} | deg(v) = k\right)$$

for any $k \in \mathbb{N}$ and any law $\mathcal{W}$ and show it is equal to :

$$\frac{1 - (1 + \frac{\ln(\beta)}{c})^k}{c},$$

where

$$\beta = \underline{\gamma} + \overline{\gamma} + \underline{\gamma}\overline{\gamma} - 1$$

with $\underline{\gamma}$ (resp. $\overline{\gamma}$) the smallest (resp. biggest) solution to $e^{-ce^{-cx}} = x$.