MAXIMUM STATIONARY VALUES IN DIRECTED RANDOM GRAPHS

Guillem Perarnau - UPC (Barcelona)

Oxford Discrete Mathematics and Probability Online Seminar

27/04/21

joint work with Xing Shi Cai, Pietro Caputo and Matteo Quattropani.

Directed configuration model (DCM)

$n \in \mathbb{N}$

$$D_n = (D_n^-, D_n^+) = ((d_1^-, d_1^+), \ldots, (d_n^-, d_n^+)) \, ,$$

heads tails

$$\sum_{i=1}^{n} d_i^- = \sum_{i=1}^{n} d_i^+ = m$$

Directed configuration model

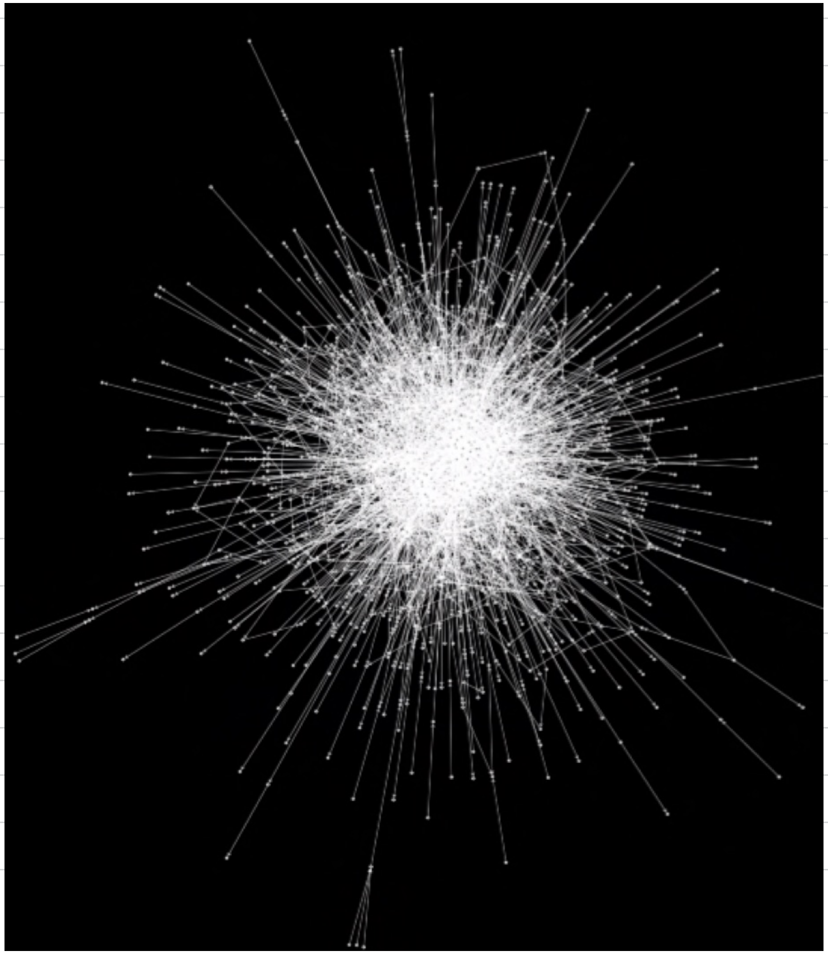
$n \in \mathbb{N}$

$$D_n = (D_n^-, D_n^+) = ((d_1^-, d_1^+), \dots, (d_n^-, d_n^+)),$$

heads tails

$$\sum_{i=1}^n d_i^- = \sum_{i=1}^n d_i^+ = m$$

$\vec{G}(D_n)$: uniform random pairing between heads and tails



Random walks on digraphs: (X_t) random walk on $\vec{G}$

$$\mathbb{P}(X_t = v \mid X_{t-1} = u) = p_{u,v} = \frac{m(u,v)}{d_u^+}, \quad m(u,v) = \#\underset{u \to v}{\text{arcs}}$$

stationary distribution: $\pi = \pi P$, $P = (p_{uv})_{u,v \in [n]}$

strongly connected component (SCC): maximal subgraph H
$\forall u,v \in V(H)$, $u \rightleftarrows v$

Random walks on digraphs: (X_t) random walk on $\vec{G}$

$$\mathbb{P}(X_t = v \mid X_{t-1} = u) = p_{u,v} = \frac{m(u,v)}{d_u^+}, \quad m(u,v) = \#\underset{u \to v}{\text{arcs}}$$

stationary distribution: $\pi = \pi P$, $P = (p_{uv})_{u,v \in [n]}$

strongly connected component (SCC): maximal subgraph H $\forall u,v \in V(H)$, $u \rightleftarrows v$

Ergodic Theorem: If $\vec{G}$ has unique attractive SCC:

i) π is unique and supported on SCC

ii) $\forall u,v \in [n]$, $\frac{1}{t} \sum_{s=1}^{t} \mathbb{P}(X_t = v \mid X_0 = u) \xrightarrow[t \to \infty]{} \pi(v)$

Random walks on random digraphs

RW on $\vec{G}(D_n)$: 2 levels of randomness $\begin{cases} \text{1 Sample } \vec{G}(D_n) \\ \text{2 Perform RW} \end{cases}$

Lemma: if $\min_v d_v^+ \geq 2$, whp unique attractive SCC ($\cap$ unique)

Random walks on random digraphs

RW on $\vec{G}(D_n)$: 2 levels of randomness $\begin{cases} ① \text{ Sample } \vec{G}(D_n) \\ ② \text{ Perform RW} \end{cases}$

$\boxed{\text{Lemma: if } \min_v d_v^+ \geq 2 \text{, whp unique attractive SCC } (\sqcap \text{ unique})}$

Assumption : $\exists \eta, C, K > 0, \forall n \in \mathbb{N}$

i) $\min_v d_v^+ \geq 2$

ii) $\max_v d_v^+ \leq K$

iii) $\sum_v (d_v^-)^{2+\eta} < C \cdot n$

Stationary distribution on $\vec{G}(D_n)$

- Undirected : $(d_1, \dots, d_n)$, $G(D_n) \Rightarrow \pi(v) = \frac{d_v}{m}$ deterministic.
- Directed : π random measure, depends on the graph geometry.

$$\mu_{in}(v) = \frac{d_v^-}{m}$$

Typical stationary values

empirical distribution: $\Psi_n = \frac{1}{n} \sum_v \delta_{m_{\Pi(v)}}$

Bordenave, Caputo, Salez '18, CCPQ '21+:

there exists $(\mathcal{L}_n)_{n \geq 1}$ deterministic laws, whp

$$\mathcal{W}_1(\Psi_n, \mathcal{L}_n) \longrightarrow 0$$

1-Wasserstein distance

$$\mathcal{W}_1(\mu, \nu) = \sup \left\{ \left| \int f \, d\mu - \int f \, d\nu \right| : f \text{ 1-Lipschitz} \right\}$$

Typical stationary values

empirical distribution: $\Psi_n = \frac{1}{n} \sum_{v} \delta_{m\pi(v)}$

Bordenave, Caputo, Salez '18, CCPQ '21+

there exists $(\mathcal{L}_n)_{n\geq 1}$ deterministic laws, whp

$$\mathcal{W}_1(\Psi_n, \mathcal{L}_n) \longrightarrow 0$$

- captures the typical values of π
- $\| \pi - \mu_{in} P^h \|_{TV} \longrightarrow 0$ for $h = c(\eta) \cdot \log n$

Extremal stationary values

- $\Pi_{max} = \max_{v} \pi(v)$
- upper tail of Ψ_n: $\Psi(n^a, \infty) = \frac{1}{n} \sum_{v \in [n]} \mathbb{1}(n\pi(v) \geq n^a)$

Motivation: π measure of popularity $\Rightarrow$ ranking (Pagerank)

Extremal stationary values

- $\Pi_{max} = \max_v \pi(v)$

- upper tail of Ψ_n: $\Psi(n^a, \infty) = \frac{1}{n} \sum_{v \in [n]} \mathbb{1}(m\pi(v) \geq n^a)$

Motivation: π measure of popularity $\Rightarrow$ ranking (Pagerank)

Remark: for π_{min} and the behaviour of the lower tail,

see Caputo and Quattropani (2020) and Cai, P. (2020+)

Π_{max} : The bounded case

Caputo and Quattropani '20: if $\Delta^- = O(1)$, whp

$$m \cdot \Pi_{max} = O\left(\log^{1-\kappa_0} n\right), \quad \text{where} \quad \kappa_0 = \frac{\log \delta^+}{\log \Delta^-} \leq 1$$

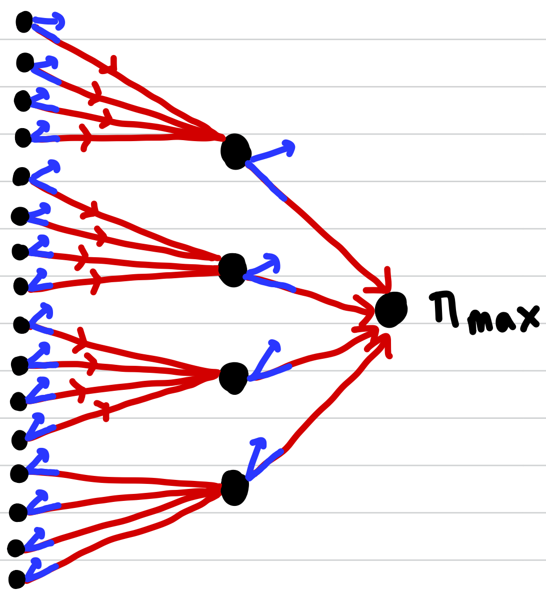
π_{max} : The bounded case

Caputo and Quattropani '20: if $\Delta^- = O(1)$, whp

$$m \cdot \pi_{max} = O\left(\log^{1-\kappa_0} n\right), \quad \text{where} \quad \kappa_0 = \frac{\log \delta^+}{\log \Delta^-} < 1$$

$\Delta^- = 4$:
many paths

$\delta^+ = 2$:
easy to follow



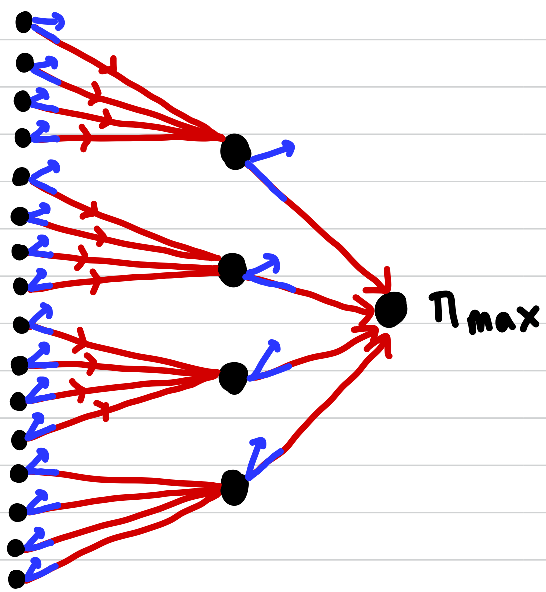
π_{max} : The bounded case

Caputo and Quattropani '20: if $\Delta^- = O(1)$, whp

$$m \cdot \pi_{max} = O\left(\log^{1-\kappa_0} n\right), \quad \text{where} \quad \kappa_0 = \frac{\log \delta^+}{\log \Delta^-} < 1$$

$\Delta^- = 4$: many paths

$\delta^+ = 2$: easy to follow



- π_{max} has logarithmic deviation with respect to

$$\max_v \mu_{in}(v) = \frac{\Delta^-}{m} = \Theta\left(\frac{1}{m}\right)$$

- If linearly many degrees (Δ^-, δ^+) then $m\pi_{max} = \Theta\left(\log^{1-\kappa_0} n\right)$

Π_{max} : the unbounded case

CCPQ 21+ : whp

$$(1+o(1))\,\Delta^{-} \leq m\cdot\Pi_{max} = O\left(\log n\cdot\Delta^{-}\right)$$

Π_{max} : the unbounded case

CCPQ 21+ : whp

$$(1+o(1))\,\Delta^- \leq m \cdot \Pi_{max} = O(\log n \cdot \Delta^-)$$

Tightness of the bounds:

- CQ 20 : examples with $m\,\pi_{max} = (\log n)^{1-o(1)}\,\Delta^-$
- CCPQ 21+: if max in-deg vertex is "outstanding", then

$$m \cdot \Pi_{max} = (1+o(1))\,\Delta^-$$

Ideas of proof for UB : $\pi_{max} \leq C \cdot (\log n)^7 \frac{\Delta^-}{m} = \pi_*$

Ideas of proof for UB : $\pi_{max} \leq C \cdot (\log n)^{7} \frac{\Delta^{-}}{m} = \pi_{*}$

$$\mu_t(y) = \mathbb{P}(X_t = y \mid X_0 \text{ u.a.r})$$

(1) For $t = \log^3 n$, $\|\pi - \mu_t\|_{TV} = o(n^{-3})$

Ideas of proof for UB : $\pi_{max} \leq C \cdot (\log n)^{7} \frac{\Delta^{-}}{m} = \pi_*$

$$\mu_t(y) = \mathbb{P}(X_t = y \mid X_0 \text{ u.a.r})$$

① For $t = \log^3 n$, $\|\pi - \mu_t\|_{TV} = o(n^{-3})$

② For $K = \frac{c}{8} \log n$, $\mathbb{E}(\mu_t^K(y)) \leq \left(\frac{2Kt^2\Delta^{-}}{m}\right)^K$

Ideas of proof for UB : $\pi_{max} \leq C \cdot (\log n)^7 \frac{\Delta^-}{m} = \pi_*$

$$\mu_t(y) = \mathbb{P}(X_t = y \mid X_0 \text{ u.a.r.})$$

① For $t = \log^3 n$, $\| \pi - \mu_t \|_{TV} = o(n^{-3})$

② For $K = \frac{c}{8} \log n$, $\mathbb{E}(\mu_t^k(y)) \leq \left(\frac{2 k t^2 \Delta^-}{m} \right)^k$

$(X_s^{(\ell)})_{s \geq 0}^{\ell \in [k]}$ independent RWs, $X_0^{(\ell)}$ indep & u.a.r., same environment

$$\mathbb{E}(\mu_t^k(y)) = \mathbb{P}(X_t^{(1)} = y, \dots, X_t^{(k)} = y)$$

Ideas of proof for UB : $\pi_{max} \leq C \cdot (\log n)^{7} \frac{\Delta^-}{m} = \pi_*$

$$\mu_t(y) = \frac{1}{n} \sum_x P^t(x,y)$$

1. For $t = \log^3 n$, $\| \pi - \mu_t \|_{TV} = o(n^{-3})$

2. For $K = \frac{c}{8} \log n$, $\mathbb{E}(\mu_t^K(y)) \leq \left(\frac{2 K t^2 \Delta^-}{m} \right)^K$

3. Markov : $\mathbb{P}(\mu_t(y) > \pi_*/2) \leq \frac{\mathbb{E}(\mu_t^K(y))}{(\pi_*/2)^K} \overset{2}{=} o\left(\frac{1}{n}\right)$

4. Union bound over all vertices, whp

$$\pi_{max} \overset{1}{\leq} \max_y \mu_t(y) + o(n^{-3}) \overset{3}{\leq} \pi_*$$

The upper tail and the power-law hypothesis (PLH)

An empirical distribution $\mu_n = \frac{1}{n}\sum_v \delta_{x_v}$ has power-law behaviour with index κ if $\forall a \in (0, 1/\kappa)$

$$\mu_n(n^a, \infty) = \frac{1}{n}\sum_v \mathbb{1}(x_v > n^a) = n^{-a\kappa + o(1)}$$

The upper tail and the power-law hypothesis (PLH)

An empirical distribution $\mu_n = \frac{1}{n} \sum_v \delta_{x_v}$ has power-law behaviour with index κ if $\forall a \in (0, 1/\kappa)$

$$\mu_n(n^a, \infty) = \frac{1}{n} \sum_v \mathbb{1}(x_v > n^a) = n^{-a\kappa + o(1)}$$

In-degrees: $\phi(n^a, \infty) = \frac{1}{n} \sum_v \mathbb{1}(d_v^- > n^a)$

Stationary: $\psi(n^a, \infty) = \frac{1}{n} \sum_v \mathbb{1}(m\,\pi(v) > n^a)$

PLH (for RWs): ϕ_n PL index $\kappa \implies \psi_n$ PL index κ

Supporting the PLH:

- Converge in 1-Wasserstein sense
  (BCS '18, CCPQ '21+)
- Upper tail properties of α_n

$\Rightarrow$ limiting distr. of Ψ_n has power-law tail

(Chen, Litvak, Olvera-Cravioto '17
Oc '12 & '19+; Volkovich, Litvak '10)

Supporting the PLH:

- Converge in 1-Wasserstein sense (BCS '18, CCPQ '21+)
- Upper tail properties of α_n (Chen, Litvak, Olvera-Cravioto '17; Oc '12 & '19+; Volkovich, Litvak '10)

$\Longrightarrow$ limiting distr. of Ψ_n has power-law tail

$$\boxed{\text{CCPQ 21+: if in-degrees } \Phi_n \text{ PL index } \kappa > 2, \text{ whp stationary values } \Psi_n \text{ PL index } \kappa.}$$

Obs: $\Psi(n^a, \infty) \lesssim n^{-a\kappa + o(1)}$ for sequences with κ-bounded moment

Pagerank and PLH

Pagerank RW: $\alpha \in [0,1]$, λ distr. on $[n]$, $(X_t^{\alpha,\lambda})$

$$\mathbb{P}(X_t^{\alpha,\lambda} = v \mid X_{t-1}^{\alpha,\lambda} = u) = (1-\alpha)\,\frac{m(u,v)}{d_u^+} + \alpha\,\lambda(v)$$

PLH: Vast amount of empirical observations,
growing mathematical evidence (CLOc '17, Oc '19)

Pagerank and PLH

Pagerank RW: $\alpha \in [0,1]$, λ distr. on $[n]$, $(X_t^{\alpha,\lambda})$

$$\mathbb{P}(X_t^{\alpha,\lambda} = v \mid X_{t-1}^{\alpha,\lambda} = u) = (1-\alpha)\frac{m(u,v)}{d_u^+} + \alpha\lambda(v)$$

PLH: Vast amount of empirical observations, growing mathematical evidence (CLOc '17, Oc '19)

$$\boxed{\begin{array}{l} \text{CCPQ 21+: in Pagerank walk, } \alpha_n < 1,\ \lambda_n \text{ "almost unif"} \\ \qquad \phi_n \text{ PL index } k > 2 \Longrightarrow \psi_n^{\alpha,\lambda} \text{ PL index } k \end{array}}$$

Open problems: Conditions: $2 \underset{\text{i)}}{\leq} d_v^+ \underset{\text{ii)}}{=} O(1)$, $\sum_v (d_v^-)^{2+\eta} \underset{\text{iii)}}{\leq} O(n)$

A) CCPQ '21+: $(1+o(1))\Delta^- \lesssim m\,\pi_{max} = O(\log n \cdot \Delta^-)$

B) CCPQ '21+: PLH for RW and Pagerank (upper tail)

Open problems: Conditions: $2 \overset{i)}{\le} d_v^+ \overset{ii)}{=} O(1)$, $\sum_v (d_v^-)^{2+\eta} \overset{iii)}{=} O(n)$

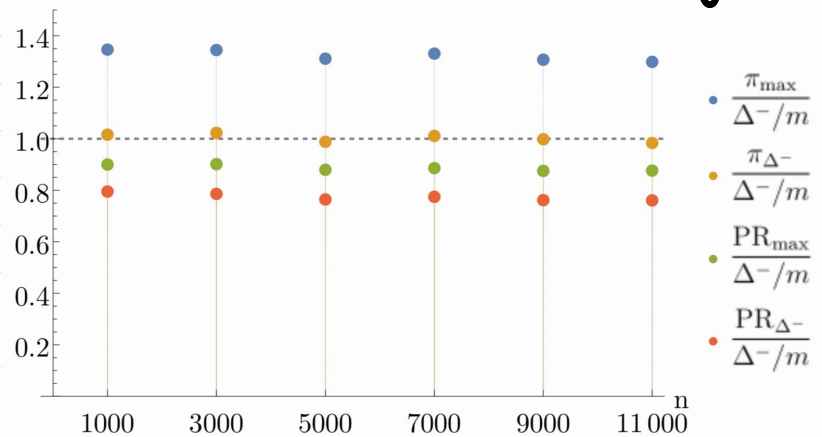
(A) CCPQ '21+: $(1+o(1))\Delta^- \le m\,\pi_{\max} = O(\log n \cdot \Delta^-)$

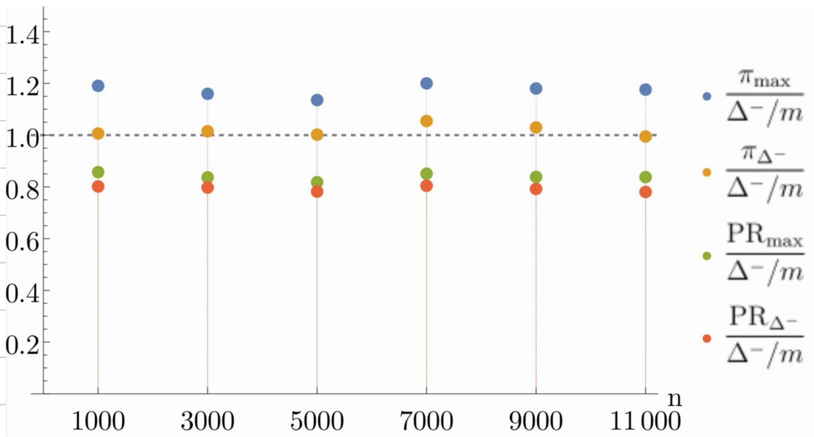
(B) CCPQ '21+: PLH for RW and Pagerank (upper tail)

P1) additional conditions to close the gap in (A).

Simulations: in-deg PL index K and $d_v^+ = 2$

$k = 5/2$





$k = 7/2$

Guess: $m\,\pi_{\max} = (c(K)+o(1))\Delta^-$, $\quad c(K) > 1$.

Open problems: Conditions: $\underset{\text{i)}}{2 \leq} d_v^+ \underset{\text{ii)}}{=} O(1)$, $\sum_v (d_v^-)^{2+\eta} \underset{\text{iii)}}{=} O(n)$

A CCPQ '21+: $(1+o(1))\Delta^- \leq m\,\pi_{max} = O(\log n \cdot \Delta^-)$

B CCPQ '21+: PLH for RW and Pagerank (upper tail)

P1) additional conditions to close the gap in A.

P2) drop ii) and, only for Pagerank, i).

Open problems: Conditions: $\underset{i)}{2 \leq} d_v^+ \underset{ii)}{=} O(1)$, $\sum_v (d_v^-)^{2+\eta} \underset{iii)}{=} O(n)$

Ⓐ CCPQ '21+: $(1+o(1))\Delta^- \leq m\,\pi_{max} = O(\log n \cdot \Delta^-)$

Ⓑ CCPQ '21+: PLH for RW and Pagerank (upper tail)

P1) additional conditions to close the gap in Ⓐ.

P2) drop ii) and, only for Pagerank, i).

P3) obtain estimations for upper tail Ψ_n for non-PL.

π_{min}:

Cai and P. '20+ : whp

$$m \cdot \pi_{min} = n^{-a_0 + o(1)}$$

where $a_0 \geq 0$ is a constant depending on D.

Remarks: i) lower tail: $\forall a \in (a_0, 1)$, $\exists b = b(a)$ s.t whp

$$\Psi(0, n^{-a}) = \frac{1}{n} \sum_v \mathbb{1}(m\pi(v) \in (0, n^{-a})) = n^{-b + o(1)}$$

ii) applic: cover time of $\vec{G}(D_n)$ is $n^{1+a_0+o(1)}$ whp