Preferential attachment trees built from random walks

Gábor Pete (Rényi Institute and TU Budapest)





János Engländer (University of Colorado, Boulder)



Giulio Iacobelli (Federal University of Rio de Janeiro)



Rodrigo Ribeiro (University of Denver)

There are many models of RWs on graphs that change over time, either independently of the walk (e.g., random walk on dynamical percolation clusters, by Peres, Sousi, Stauffer, Steif),

or the walk is changing the transition probabilities (e.g., reinforced random walk by Merkl-Rolles, Angel-Crawford-Kozma, ..., true self-repelling motion by Tóth-Werner).

There are many models of RWs on graphs that change over time, either independently of the walk (e.g., random walk on dynamical percolation clusters, by Peres, Sousi, Stauffer, Steif),

or the walk is changing the transition probabilities (e.g., reinforced random walk by Merkl-Rolles, Angel-Crawford-Kozma, ..., true self-repelling motion by Tóth-Werner).

Now the the walker is building the graph: Fix $\gamma \in [0, 1]$. In step *n*:

- add a leaf with probability $p_n = n^{-\gamma}$ to your current vertex X_n ,
- step to uniform random neighbor X_{n+1} on new tree T_{n+1} .

イロト イポト イヨト イヨト 二日

There are many models of RWs on graphs that change over time, either independently of the walk (e.g., random walk on dynamical percolation clusters, by Peres, Sousi, Stauffer, Steif),

or the walk is changing the transition probabilities (e.g., reinforced random walk by Merkl-Rolles, Angel-Crawford-Kozma, ..., true self-repelling motion by Tóth-Werner).

Now the the walker is building the graph: Fix $\gamma \in [0, 1]$. In step *n*:

- add a leaf with probability $p_n = n^{-\gamma}$ to your current vertex X_n ,
- step to uniform random neighbor X_{n+1} on new tree T_{n+1} .

Introduced by Amorim, Figueiredo, Iacobelli, Neglia (2016). Mixture of RW (recurrence, transience, ballisticity) and random graph questions (diameter, degree distribution).

イロト イポト イヨト イヨト 三日



(Thanks to Ágnes Kúsz for the pictures.)

◆□▶ ◆□▶ ◆ □▶ ◆ □▶ ●

5900

3

$\gamma =$ 0, always grow a leaf:

If X_n is a leaf, $\mathbb{E}[\operatorname{dist}(o, X_{n+1})] = \mathbb{E}[\operatorname{dist}(o, X_n)].$

If X_n is not a leaf, $\mathbb{E}[\operatorname{dist}(o, X_{n+1})] \ge \mathbb{E}[\operatorname{dist}(o, X_n)] + 1/3$.

 X_n spends at least a constant proportion of time at non-leaves.

・ 同 ト ・ ヨ ト ・ ヨ ト

$\gamma =$ 0, always grow a leaf:

If X_n is a leaf, $\mathbb{E}[\operatorname{dist}(o, X_{n+1})] = \mathbb{E}[\operatorname{dist}(o, X_n)].$

If X_n is not a leaf, $\mathbb{E}[\operatorname{dist}(o, X_{n+1})] \ge \mathbb{E}[\operatorname{dist}(o, X_n)] + 1/3$.

 X_n spends at least a constant proportion of time at non-leaves.

 \implies ballistic X_n , linearly growing T_n .

This and much more was proved by Figueiredo, Iacobelli, Oliveira, Reed, Ribeiro (2021), and Iacobelli, Ribeiro, Valle, Zuaznábar (2022) in the elliptic regime: $p_n > c > 0$, and versions of that.

・ 同 ト ・ ヨ ト ・ ヨ ト

How much are we walking between growth times?

3

SQA

How much are we walking between growth times?

Mixing time of a tree:

 $T_{\min} \leq \operatorname{vol}^{1+o(1)} \operatorname{diam}$, and typically $T_{\min} \geq c \operatorname{vol}$.

Also, *typical* hitting time of root from a distant leaf is $\tau_{hit} \ge c \text{ vol.}$

One stage: the walk between growth times. By time *n*, around $n^{1-\gamma}$ growth times, so $\operatorname{vol}_n \approx n^{1-\gamma}$. The length $S_{n^{1-\gamma}}$ of the $n^{1-\gamma}$ th stage is typically $\approx n^{\gamma}$.

・ロト ・ 同ト ・ ヨト ・ ヨト ・ ヨ

How much are we walking between growth times?

Mixing time of a tree:

 $T_{\text{mix}} \leq \text{vol}^{1+o(1)} \text{ diam, and typically } T_{\text{mix}} \geq c \text{ vol.}$

Also, *typical* hitting time of root from a distant leaf is $\tau_{hit} \ge c \text{ vol.}$

One stage: the walk between growth times. By time *n*, around $n^{1-\gamma}$ growth times, so $\operatorname{vol}_n \approx n^{1-\gamma}$. The length $S_{n^{1-\gamma}}$ of the $n^{1-\gamma}$ th stage is typically $\approx n^{\gamma}$.

$\gamma < 1/2$:

 $S_{n^{1-\gamma}} \approx n^{\gamma} \ll n^{1-\gamma} \leq \tau_{\text{hit}}(T_n)$, so the chance that $(X_t)_{t\geq 0}$ hits o in a stage should be much smaller than the stationary distribution $1/\text{vol}_n$, should happen only finitely often, by Borel-Cantelli.

 \implies transience should hold.

How much are we walking between growth times?

Mixing time of a tree:

 $T_{\text{mix}} < \text{vol}^{1+o(1)} \text{ diam, and typically } T_{\text{mix}} \ge c \text{ vol.}$

Also, *typical* hitting time of root from a distant leaf is $\tau_{hit} \ge c \text{ vol.}$

One stage: the walk between growth times. By time *n*, around $n^{1-\gamma}$ growth times, so $\operatorname{vol}_n \approx n^{1-\gamma}$. The length $S_{n^{1-\gamma}}$ of the $n^{1-\gamma}$ th stage is typically $\approx n^{\gamma}$.

$\gamma < 1/2$:

 $S_{n^{1-\gamma}} \approx n^{\gamma} \ll n^{1-\gamma} \leq \tau_{\text{hit}}(T_n)$, so the chance that $(X_t)_{t\geq 0}$ hits o in a stage should be much smaller than the stationary distribution $1/vol_n$, should happen only finitely often, by Borel-Cantelli.

 \implies transience should hold. Nothing is proved. $\boxed{2}$



・ロト ・ 同ト ・ ヨト ・ ヨト - ヨ

By time *n*, around $n^{1-\gamma}$ stages, so $\operatorname{vol}_n \approx n^{1-\gamma}$. The length $S_{n^{1-\gamma}}$ of the $n^{1-\gamma}$ th stage is typically $\approx n^{\gamma}$. Typically, $c \operatorname{vol} \leq T_{\text{mix}} \leq \operatorname{vol}^{1+o(1)} \operatorname{diam}$.

 $\gamma > 1/2$: Hope to quite mix in a stage, so T_n is more balanced, hence diam is smaller, hence we get more easily mixed.

・ロト ・ 雪 ト ・ ヨ ト ・

By time *n*, around $n^{1-\gamma}$ stages, so $\operatorname{vol}_n \approx n^{1-\gamma}$. The length $S_{n^{1-\gamma}}$ of the $n^{1-\gamma}$ th stage is typically $\approx n^{\gamma}$. Typically, $c \operatorname{vol} \leq T_{\mathsf{mix}} \leq \operatorname{vol}^{1+o(1)} \operatorname{diam}$.

 $\gamma > 1/2$: Hope to quite mix in a stage, so T_n is more balanced, hence diam is smaller, hence we get more easily mixed.

Key proposition. For $\gamma > 1/2$, have diam $(T_n) = \operatorname{vol}_n^{o(1)} w.h.p.$

Thus $T_{\text{mix}} = \text{vol}_n^{1+o(1)}$, so $\mathbb{P}[S_k < T_{\text{mix}}] = \mathbb{P}[S_k < k^{1+o(1)}] = k^{\frac{-\gamma}{1-\gamma}}k^{1+o(1)} = k^{\frac{1-2\gamma}{1-\gamma}+o(1)}$. For $\gamma \in (1/2, 2/3]$, this is $k^{-\varepsilon}$ with $\varepsilon \leq 1$, small, but happens i.o. For $\gamma \in (2/3, 1]$, this is summable in k, happens only fin. often.

・ロト ・ 同ト ・ ヨト ・ ヨト ・ ヨ

By time *n*, around $n^{1-\gamma}$ stages, so $\operatorname{vol}_n \approx n^{1-\gamma}$. The length $S_{n^{1-\gamma}}$ of the $n^{1-\gamma}$ th stage is typically $\approx n^{\gamma}$. Typically, $c \operatorname{vol} \leq T_{\mathsf{mix}} \leq \operatorname{vol}^{1+o(1)} \operatorname{diam}$.

 $\gamma > 1/2$: Hope to quite mix in a stage, so T_n is more balanced, hence diam is smaller, hence we get more easily mixed.

Key proposition. For $\gamma > 1/2$, have diam $(T_n) = \operatorname{vol}_n^{o(1)} w.h.p.$

Thus $T_{\text{mix}} = \operatorname{vol}_n^{1+o(1)}$, so $\mathbb{P}[S_k < T_{\text{mix}}] = \mathbb{P}[S_k < k^{1+o(1)}] = k^{\frac{-\gamma}{1-\gamma}}k^{1+o(1)} = k^{\frac{1-2\gamma}{1-\gamma}+o(1)}$. For $\gamma \in (1/2, 2/3]$, this is $k^{-\varepsilon}$ with $\varepsilon \leq 1$, small, but happens i.o. For $\gamma \in (2/3, 1]$, this is summable in k, happens only fin. often. But actual proofs? What does "mixing has happened" mean?

・ロット (空)・ (日)・ (日)・ 日

For $\gamma > 1/2$, have $\mathbb{P}[\tau_{\text{hit}} < S_k] > c > 0$, hence it happens i.o. \implies recurrence.

イロト イポト イヨト イヨト 二日

SQC

For $\gamma > 1/2$, have $\mathbb{P}[\tau_{hit} < S_k] > c > 0$, hence it happens i.o. \implies recurrence.

For $\gamma > 2/3$: want to say that { $S_k < T_{mix}$ fin. oft.} means that growth happens at a stationary vertex from some random time on, hence it is the Barabási-Albert preferential attachment tree started from a random tree.

・ 同 ト ・ ヨ ト ・ ヨ ト

For $\gamma > 1/2$, have $\mathbb{P}[\tau_{hit} < S_k] > c > 0$, hence it happens i.o. \implies recurrence.

For $\gamma > 2/3$: want to say that $\{S_k < T_{mix} \text{ fin. oft.}\}\$ means that growth happens at a stationary vertex from some random time on, hence it is the Barabási-Albert preferential attachment tree started from a random tree.

Key tool (Aldous-Diaconis 1987). For any finite Markov chain $(X_t)_{t\geq 0}$ with stationary distribution π , for any starting state x, there is an optimal strong stationary stopping time η_x : $\mathbb{P}_x[X_{\eta_x} = y \mid \eta_x = t] = \pi(y)$, and $\mathbb{P}_x[\eta_x > t] = s_x(t)$, the separation distance at time t.

With this, we can produce a coupling with BA-tree process from some random time on.

TBRW: structural corollaries $\gamma > 2/3$

All almost sure BA-tree limit results that do not depend on the starting tree (B. Pittel, T. Móri, Zs. Katona) can get transfered:

$$\frac{\operatorname{diam}(T_n)}{\log n} \to c$$

$$\frac{\operatorname{dist}(o, \operatorname{Unif}(T_n))}{\log n} \to 1/2$$

$$\frac{\operatorname{max} \operatorname{deg}(T_n)}{\sqrt{\operatorname{vol}_n}} \to \zeta \text{ with a non-trivial distribution}$$

$$\frac{|\{v \in T_n : \operatorname{deg}(v) = k\}|}{\operatorname{vol}_n} \to \frac{4}{k(k+1)(k+2)}$$

イロト イヨト イヨト

TBRW: open questions

Show transience for γ < 1/2.
 Diameter of T_n is n^{1-γ-δ(γ)} for what δ(γ)?
 Degree distribution: exponential or polynomial or in between?

② What happens at
$$\gamma = 1/2$$
?

- Solution For γ ∈ (1/2, 2/3], the process is not the BA-tree process, but do we still have BA-like statistics?
- What is the scaling limit of the height process of the random walker on a BA tree? It is typically at height $(1/2) \log n$, with Gaussian fluctuations $\approx \sqrt{\log n}$, by results of Zsolt Katona. Do we see an Ornstein-Uhlenbeck process?

・ロト ・ 雪 ト ・ ヨ ト ・

A network-of-networks model for physical networks



Ivan Bonamassa (Central European University, Vienna)



Márton Pósfai (Central European University, Vienna)



Sigurdur Örn Stefánsson (University of Iceland)



Ádám Timár (University of Iceland and Rényi Institute)

Representing physical networks

Physical network: not only is the graph embedded in space, but the vertices and edges are non-overlapping physical objects.

- Traditional: balls and tubes,
- More realistic: nodes are extended objects and links are the points of contact. E.g., neurons and synapses



Representing physical networks

Physical network: not only is the graph embedded in space, but the vertices and edges are non-overlapping physical objects.

- Traditional: balls and tubes,
- More realistic: nodes are extended objects and links are the points of contact. E.g., neurons and synapses



Nodes will be random walk paths (pieces) in some finite graph *H*. In principle one could choose simple random walk, self-avoiding walk, loop-erased walk...





Nodes will be random walk paths (pieces) in some finite graph *H*. In principle one could choose simple random walk, self-avoiding walk, loop-erased walk...



伺下 イヨト イヨト

Nodes will be random walk paths (*pieces*) in some finite graph *H*. In principle one could choose simple random walk, self-avoiding walk, loop-erased walk...



□ ► < E ► < E ►</p>

Nodes will be random walk paths (*pieces*) in some finite graph *H*. In principle one could choose simple random walk, self-avoiding walk, loop-erased walk...



Nodes will be random walk paths (*pieces*) in some finite graph *H*. In principle one could choose simple random walk, self-avoiding walk, loop-erased walk...



→ Ξ →

We will mainly focus on models generated by the loop-erased random walk (LERW).

LERW from x_1 to x_0 : run simple random walk and erase all the loops as they are created.



伺 ト イヨト イヨト

We will mainly focus on models generated by the loop-erased random walk (LERW).

LERW from x_1 to x_0 : run simple random walk and erase all the loops as they are created.



伺 ト イヨト イヨト

The resulting network-of-networks in 3 dimensions:





◆□ ▶ ◆□ ▶ ◆ □ ▶ ◆ □ ▶

Э

DQC

We will look at the fully packed model: every *H*-vertex is contained in some piece.



• = • •

MQ (P

Assume the pieces in \mathbb{Z}^d have "fractal dimension" ζ , with $2\zeta \ge d$. Box of t^{th} piece: side-length ℓ_t . Typically, $\ell_s \ge \ell_t$ for $s \le t$.



If larger piece \mathcal{V}_s intersects the box of the smaller piece \mathcal{V}_t , then, by $2\zeta \ge d$, they overlap with positive probability.

Tile the box $[L]^d$ with $(L/\ell_t)^d$ boxes of side-length ℓ_t . Number of boxes intersected by \mathcal{V}_s is $\approx (\ell_s/\ell_t)^{\zeta}$.

Thus the intersection probability, randomly placed, using $v_s = |\mathcal{V}_s| \approx \ell_s^{\zeta}$:

・ 同 ト ・ ヨ ト ・ ヨ ト

$$p_{s,t} \approx \left(\frac{\ell_s}{\ell_t}\right)^{\zeta} / \left(\frac{L}{\ell_t}\right)^d = \frac{v_s v_t^{d/\zeta - 1}}{L^d}.$$

Probability that the piece added at time t intersects any existing piece s < t is approximately

$$\sum_{s:s < t} p_{s,t} = \sum_{s:s < t} w_{t-1} v_t^{d/\zeta - 1} / L^d,$$

where $w_s = v_1 + \cdots + v_s$.

In our growth process, \mathcal{V}_t grows until it hits existing piece, until the above intersection probability becomes ≈ 1 . I.e.:

$$v_t \approx \left(w_{t-1}/L^d \right)^{-rac{\zeta}{d-\zeta}}$$

This is an ODE, $v_t = w'_t$, with $w_1 \approx L^{\zeta}$. Get $w_t \approx L^d (t/L^d)^{1-\zeta/d}$ and $v_t \approx (t/L^d)^{-\zeta/d}$, and the degree of piece t at time T is

$$\deg_{\mathcal{T}}(t) \approx 1 + \frac{v_t}{L^d} \sum_{s=t}^{T} v_s^{d/\zeta-1} \approx (T/t)^{\zeta/d}.$$

It follows that degree distribution is $\mathbb{P}\left[\deg_T(\sigma_T) > k\right] \approx k^{-d/\zeta}$.

For $2\zeta < d$, get $\mathbb{P}\left[\deg_{\mathcal{T}}(\sigma_{\mathcal{T}}) > k\right] \approx k^{-2}$, mean-field, as in BA.

Example 1. LERW in $d \ge 5$.

DQ P

For $2\zeta < d$, get $\mathbb{P}\left[\deg_{\mathcal{T}}(\sigma_{\mathcal{T}}) > k\right] \approx k^{-2}$, mean-field, as in BA.

Example 1. LERW in $d \ge 5$.

Example 2. $\zeta = 1$, d = 2, random ray model:



Ex1 is really a version of the BA-tree, not just in deg-distribution.

500

For the case of LERW, mathematically rigorous proofs can be given, using connection with the Uniform Spanning Tree:

Theorem

The degree distribution of the abstract network satisfies

- $\mathbb{P}[\deg(\sigma) > t] = t^{-8/5 + o(1)}$ for dim= 2;
- $\mathbb{P}[\deg(\sigma) > t] = t^{-2+o(1)}$ for dim ≥ 5 .

Consider a Uniform Spanning Tree (UST) T in G (finite). Take a uniform random ordering x_0, x_1, \ldots, x_N of the vertices.

$$P_1 :=$$
 the path from x_1 to x_0 in T
 $P_k :=$ the path from x_k to $P_1 \cup \ldots \cup P_{k-1}$



Consider a Uniform Spanning Tree (UST) T in G (finite). Take a uniform random ordering x_0, x_1, \ldots, x_N of the vertices.

$$P_1 :=$$
 the path from x_1 to x_0 in T
 $P_k :=$ the path from x_k to $P_1 \cup \ldots \cup P_{k-1}$



Consider a Uniform Spanning Tree (UST) T in G (finite). Take a uniform random ordering x_0, x_1, \ldots, x_N of the vertices.

$$P_1 :=$$
 the path from x_1 to x_0 in T
 $P_k :=$ the path from x_k to $P_1 \cup \ldots \cup P_{k-1}$



Consider a Uniform Spanning Tree (UST) T in G (finite). Take a uniform random ordering x_0, x_1, \ldots, x_N of the vertices.

$$P_1 :=$$
 the path from x_1 to x_0 in T
 $P_k :=$ the path from x_k to $P_1 \cup \ldots \cup P_{k-1}$



Consider a Uniform Spanning Tree (UST) T in G (finite). Take a uniform random ordering x_0, x_1, \ldots, x_N of the vertices.

$$P_1 :=$$
 the path from x_1 to x_0 in T
 $P_k :=$ the path from x_k to $P_1 \cup \ldots \cup P_{k-1}$



Consider a Uniform Spanning Tree (UST) T in G (finite). Take a uniform random ordering x_0, x_1, \ldots, x_N of the vertices.

$$P_1 :=$$
 the path from x_1 to x_0 in T
 $P_k :=$ the path from x_k to $P_1 \cup \ldots \cup P_{k-1}$



One consequence of the alternative description is the existence of the infinite-volume limit.

Theorem

Let *G* be an infinite transitive graph and G_n be an exhaustion by finite induced subgraphs. The LERW-generated physical network on G_n has a weak limit, invariant under the automorphisms of *G*. The abstract network corresponding to the weak limit is a unimodular random forest consisting of one-ended trees whenever *G* has superlinear growth.

伺 ト イヨト イヨト

One consequence of the alternative description is the existence of the infinite-volume limit.

Theorem

Let *G* be an infinite transitive graph and G_n be an exhaustion by finite induced subgraphs. The LERW-generated physical network on G_n has a weak limit, invariant under the automorphisms of *G*. The abstract network corresponding to the weak limit is a unimodular random forest consisting of one-ended trees whenever *G* has superlinear growth.

One can also construct a scaling limit, using work of Archer, Nachmias, Shalev (2021) for $d \ge 5$. (In progress.)

Gives answers to questions like: *how does the subtree induced by the first 100 pieces in the construction look like*?

< 回 > < 回 > < 回 >

Degree exponent, rough sketch of proof



Want: degree of a uniformly randomly selected node of the abstract graph.

∃ > ∢

Degree exponent, rough sketch of proof



Want: degree of a uniformly randomly selected node of the abstract graph.

• = • •



This is the same as the the "degree" of the piece of a random vertex x in H. We choose x with a bias $|P(x)|^{-1}$ where P(x) is its piece.

A ►

→ Ξ →

Degree exponent, rough sketch of proof



Lemma 1: Degree of a piece is of the same magnitude as its length.

• = • •



Lemma 2: Length of P(x) is comparable to the diameter of the subtree of vertices in the UST that x separates from x_0 (the *past* of x).

So we need the distribution of this latter diameter, $\operatorname{diam}(\operatorname{past}_x)$.

Degree exponent, rough sketch of proof

Dimension 2:

growth exponent is 5/4, Kenyon (2000), Barlow-Masson (2011) Ignoring o(1) corrections in the exponents,

 $\mathbb{P}[\operatorname{diam}(\operatorname{past}_x) > r] \approx \mathbb{P}[\operatorname{Eucl-diam}(\operatorname{past}_x)^{5/4} > r]$

 $= \mathbb{P}[\text{Eucl-diam}(\text{past}_x) > r^{4/5}] \approx (r^{4/5})^{-3/4} = r^{-3/5},$

where the last \approx is by a result of Masson about intersecting LERW and SRW:



白マイド・トレ

Degree exponent, rough sketch of proof

Thus, for the uniformly selected point *o* of the *abstract network* we can summarize:

$$\mathbb{P}[\deg(o) = k] \approx \frac{1}{k} \mathbb{P}[|P(x)| = k] \approx \frac{1}{k} \mathbb{P}[\operatorname{diam}(\operatorname{past}_{x}) = k]$$
$$\approx \frac{1}{k} k^{-3/5-1} = k^{-2.6}.$$

Dimension \geq 5:

Follows from the known distribution of Eucl-diam($past_x$), Bhupatiraju-Hanson-Járai (2017), and the concentration of diam($past_x$) around Eucl-diam($past_x$)², by Lawler (1980).

・ 同 ト ・ ヨ ト ・ ヨ ト

Combinatorial Laplacian



Physical Laplacian



.Traditional approach of network science: look at $\mathbf{Q}_{\mathcal{G}}$

.What is the role of the physicality layout?

.Setup:

-Coupling strength inside a physical node = 1

-

nar

-Between nodes = w

-Small w limit

- . Shape of ~25,000 neurons + location of synapses
- . Data published on several levels
- Raw images: >20 TB
- Skeletonized neurons: >117 million segments
- Connectome adjacency matrix + neuron volumes





- **9** Show $\gamma = 3$ degree distribution for random ray model in 2d.
- Output of the Understand how exactly the Physical Laplacian feels physicality.

For instance, prove that the spectrum does not have a fat tail.

・ 同 ト ・ ヨ ト ・ ヨ ト

Thank you for your attention!

ヘロト ヘロト ヘビト ヘビト

E

990