Scaling limit of critical random trees in random environment

with Daniel Kious
Guillaume Conchon-Kerjan

① Scaling limit of a "classical" Galton Watson tree.

② Scaling limit of a Galton Watson tree in "random environment".

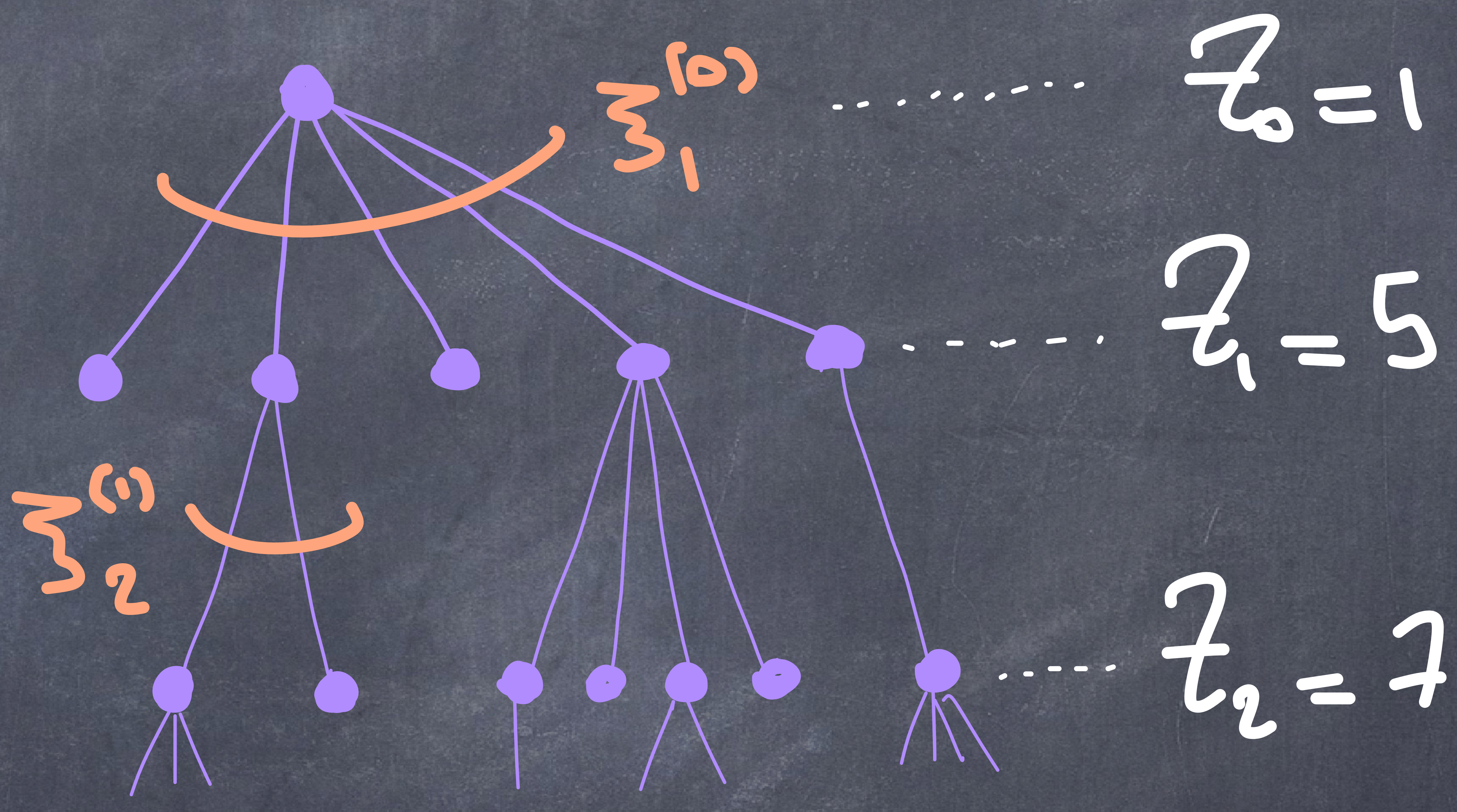
What is a GW tree?

Z_n = size of a population at generation n

$Z_0 = 1$

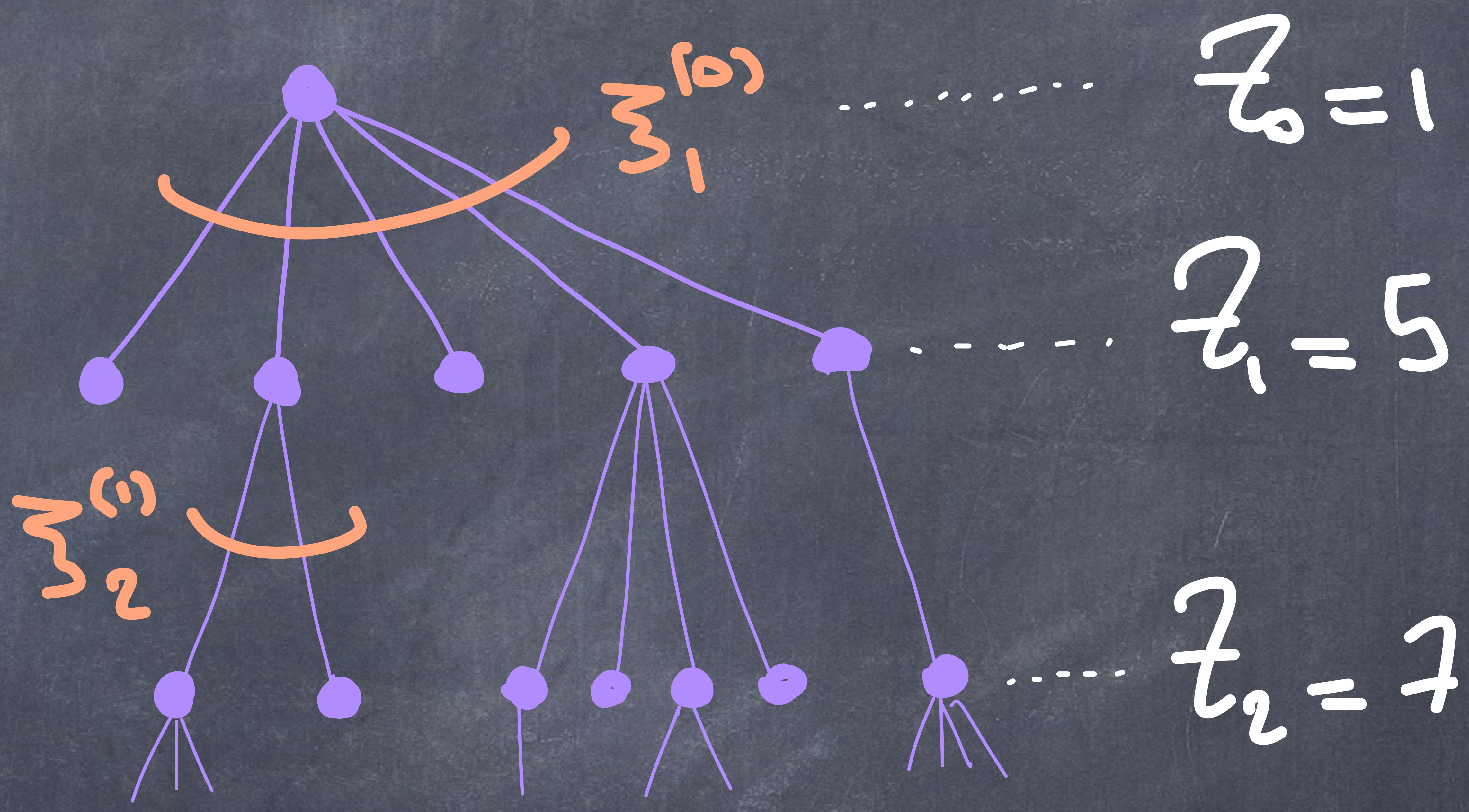
$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_i^{(n)}$$

where $\left(\xi_i^{(n)}\right)_{\substack{n \geq 0 \\ i \geq 1}}$ are iid of distribution μ on $\mathbb{N}$



What is a GW tree?

Let $\bar{\mu} = \mathbb{E}[\xi_1^{(0)}]$
(average # of offspring)



If $\bar{\mu} \le 1$ then $Z_n \xrightarrow[n\uparrow\infty]{a.s.} 0$

(sub) critical

If $\bar{\mu} > 1$ then

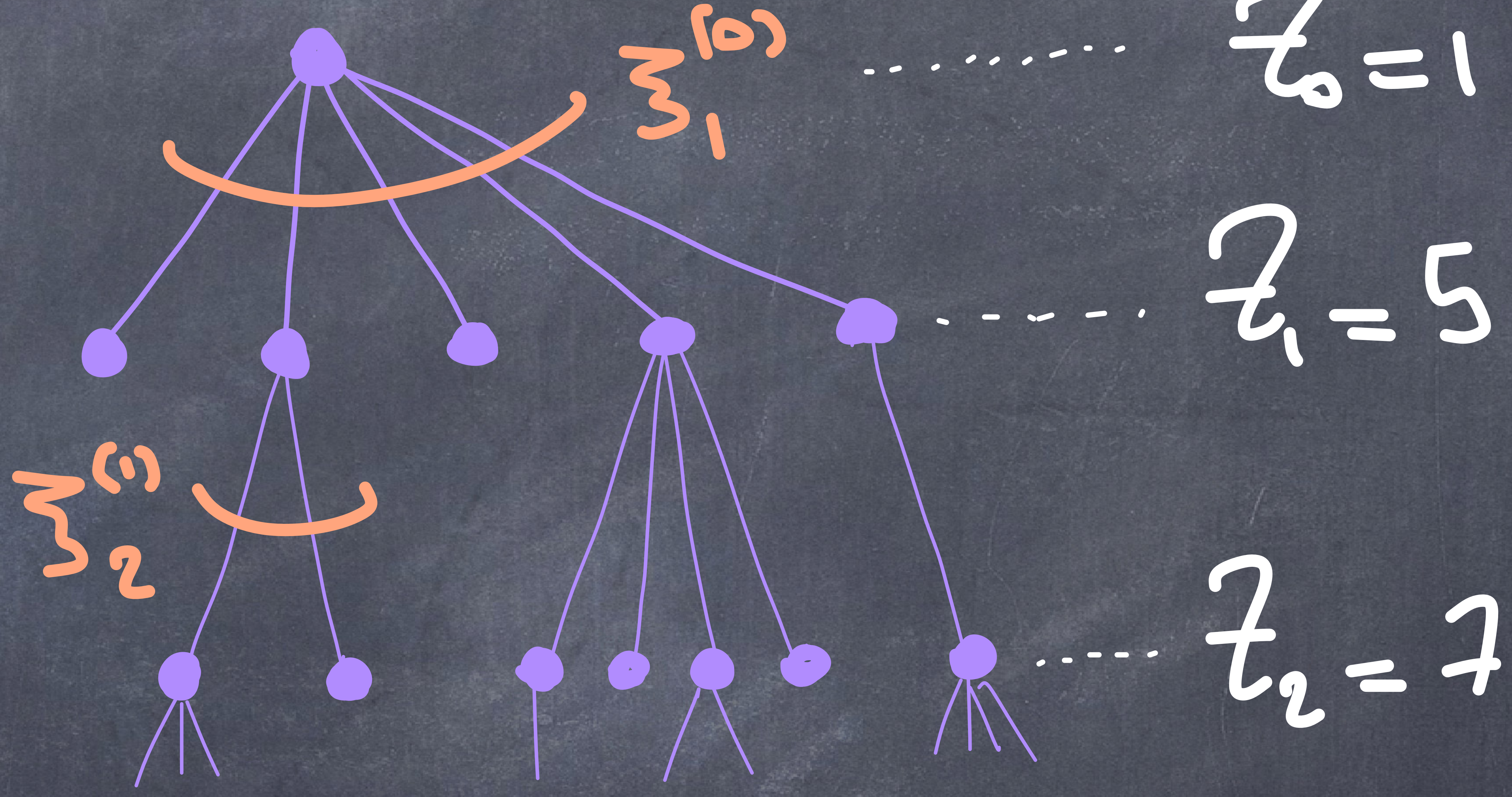
$\mathbb{P}(Z_n \longrightarrow +\infty) > 0$

super-critical

What is a GW tree?

Let $\bar{\mu} = \mathbb{E}[\xi_1^{(0)}]$

(average # of offspring)



If $\bar{\mu} \leq 1$ then $Z_n \xrightarrow[n\uparrow\infty]{a.s.} 0$

(sub) critical

If $\bar{\mu} > 1$ then

$\mathbb{P}(Z_n \longrightarrow +\infty) > 0$

super - critical

We are interested in the "critical case": $\bar{\mu} = 1$.

Scaling limit of the classical GW tree

Let $\mathcal{T}^{(n)}$ be the GW tree of offspring distribution μ conditioned on its total population to be at least n

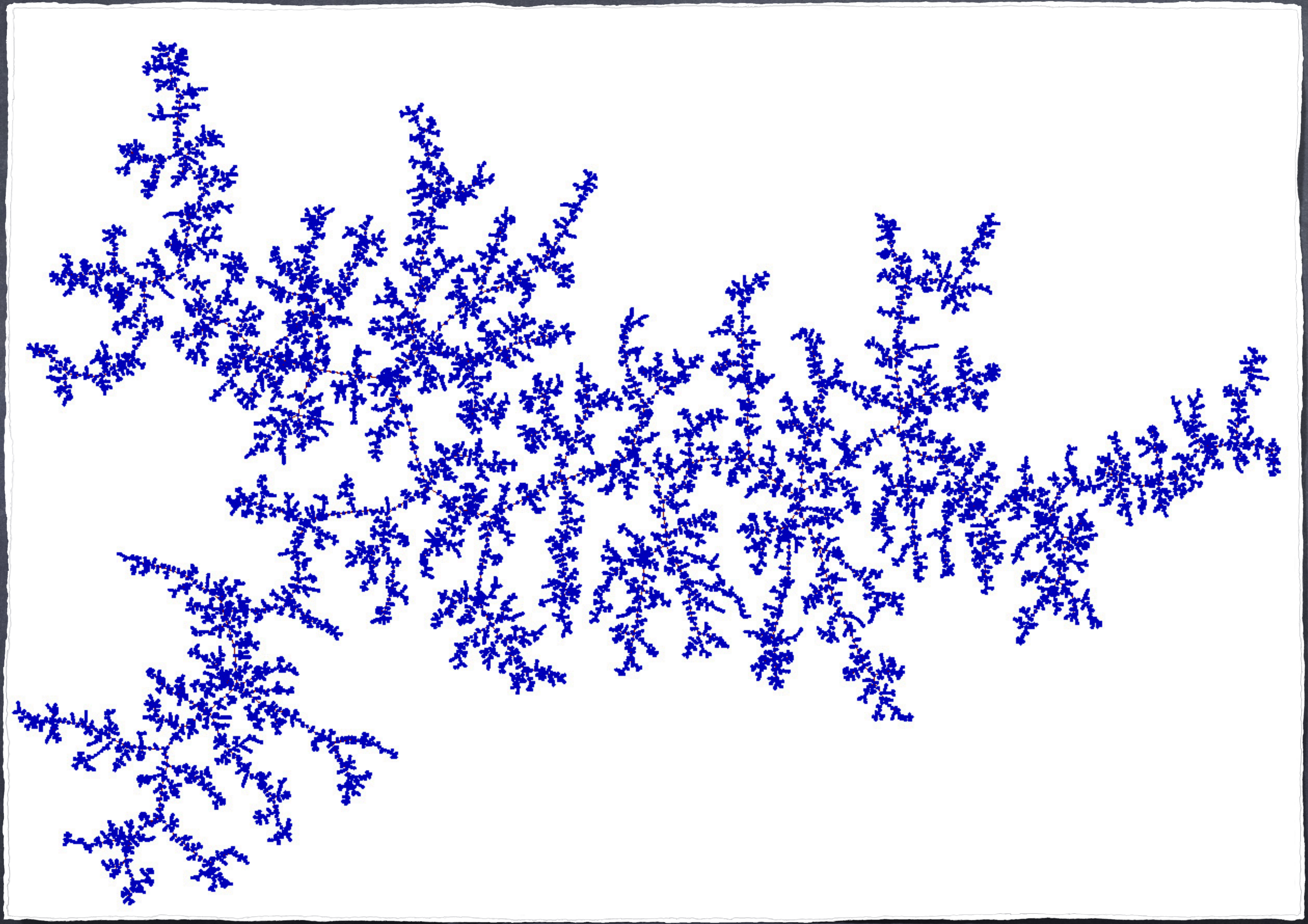
Th (Aldous '91): if $\sigma^2 := \mathrm{Var}(\mu) < \infty$, then

$$\frac{1}{\sigma\sqrt{n}}\mathcal{T}^{(n)} \xrightarrow[n\uparrow\infty]{(d)} \mathrm{CRT}$$

for the Gromov-Haussdorff (GH) metric on compact metric spaces.

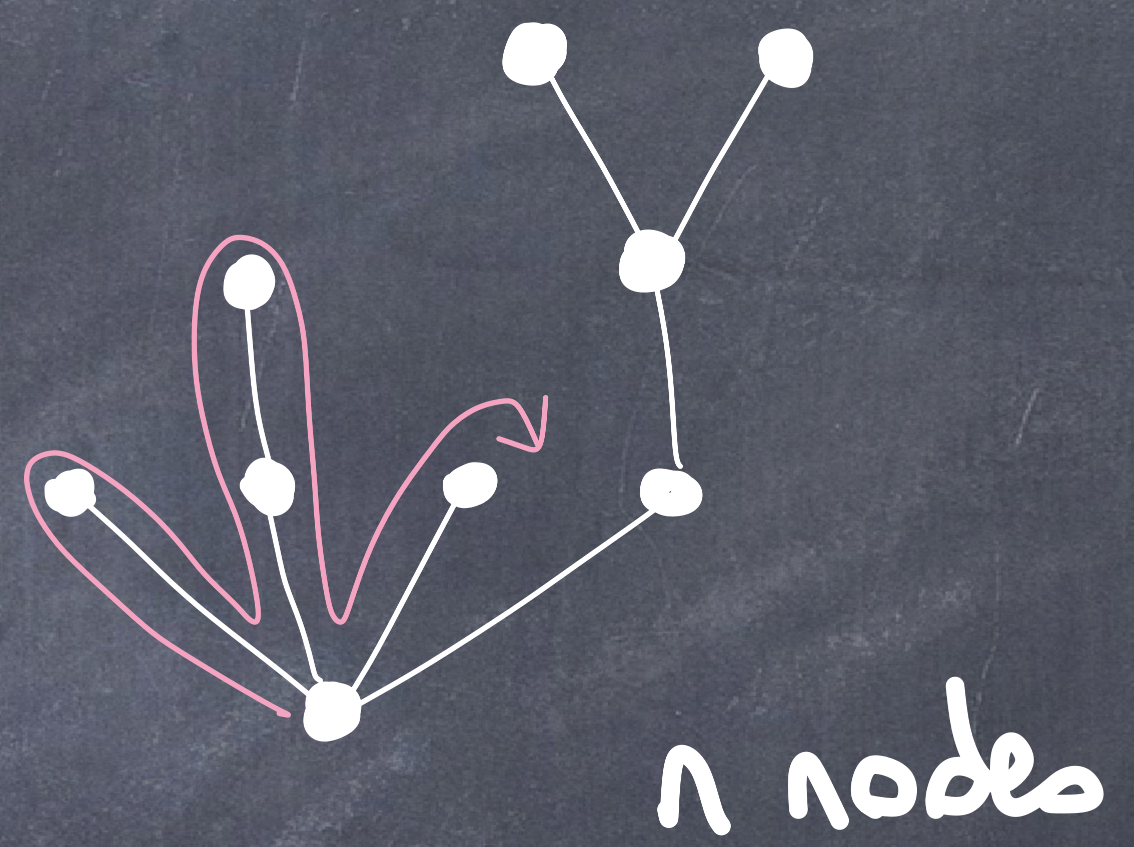
seen as a compact metric space.

Aldous' continuum random tree.



Trees as metric spaces.

A tree is a metric space:



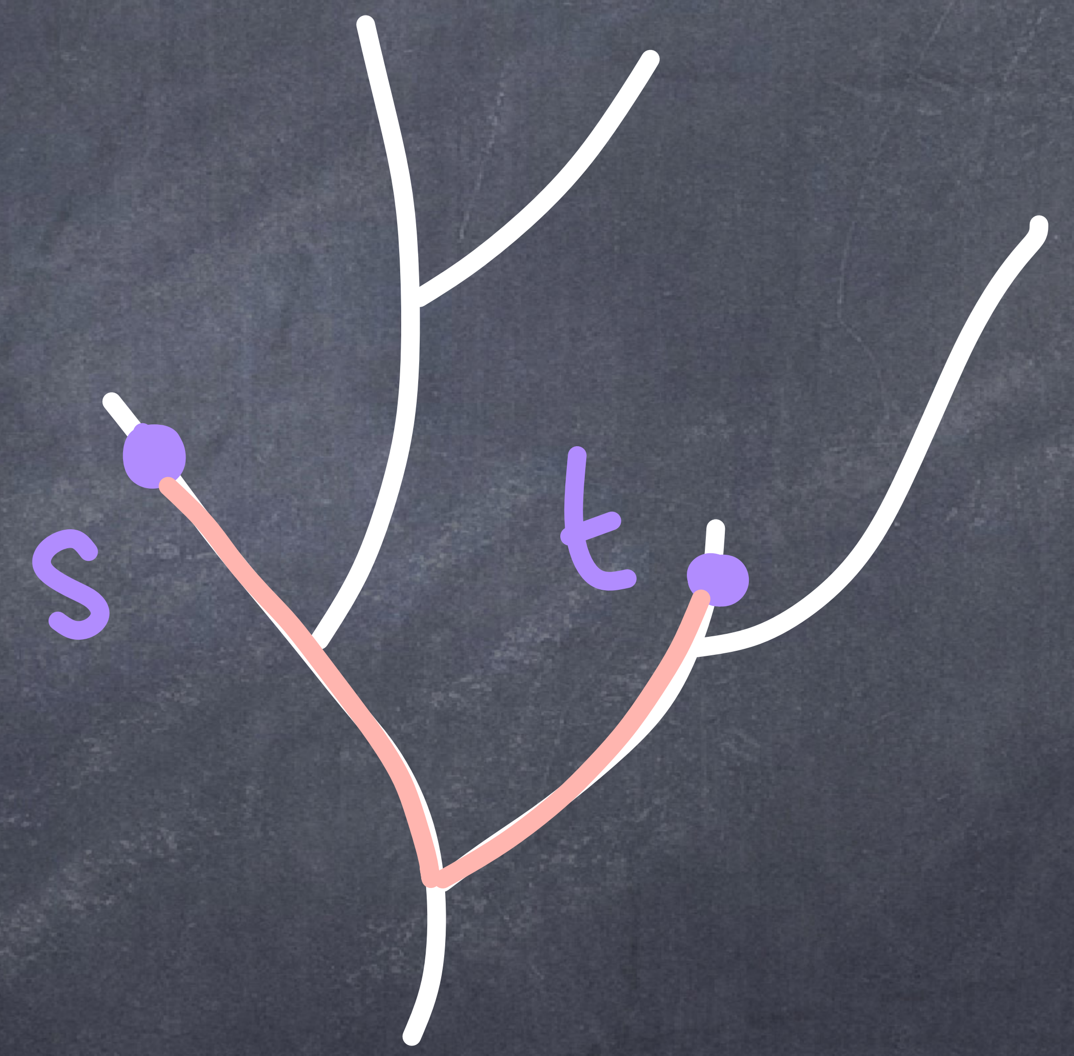
excursions in the upper half-plane = trees

We call this the "contour function" of the tree

Trees as metric spaces.

A tree is a metric space: $\forall\, s,t \in [0,L]$

$$d(s,t) = f(s) + f(t) - 2 \min_{[s,t]} f$$

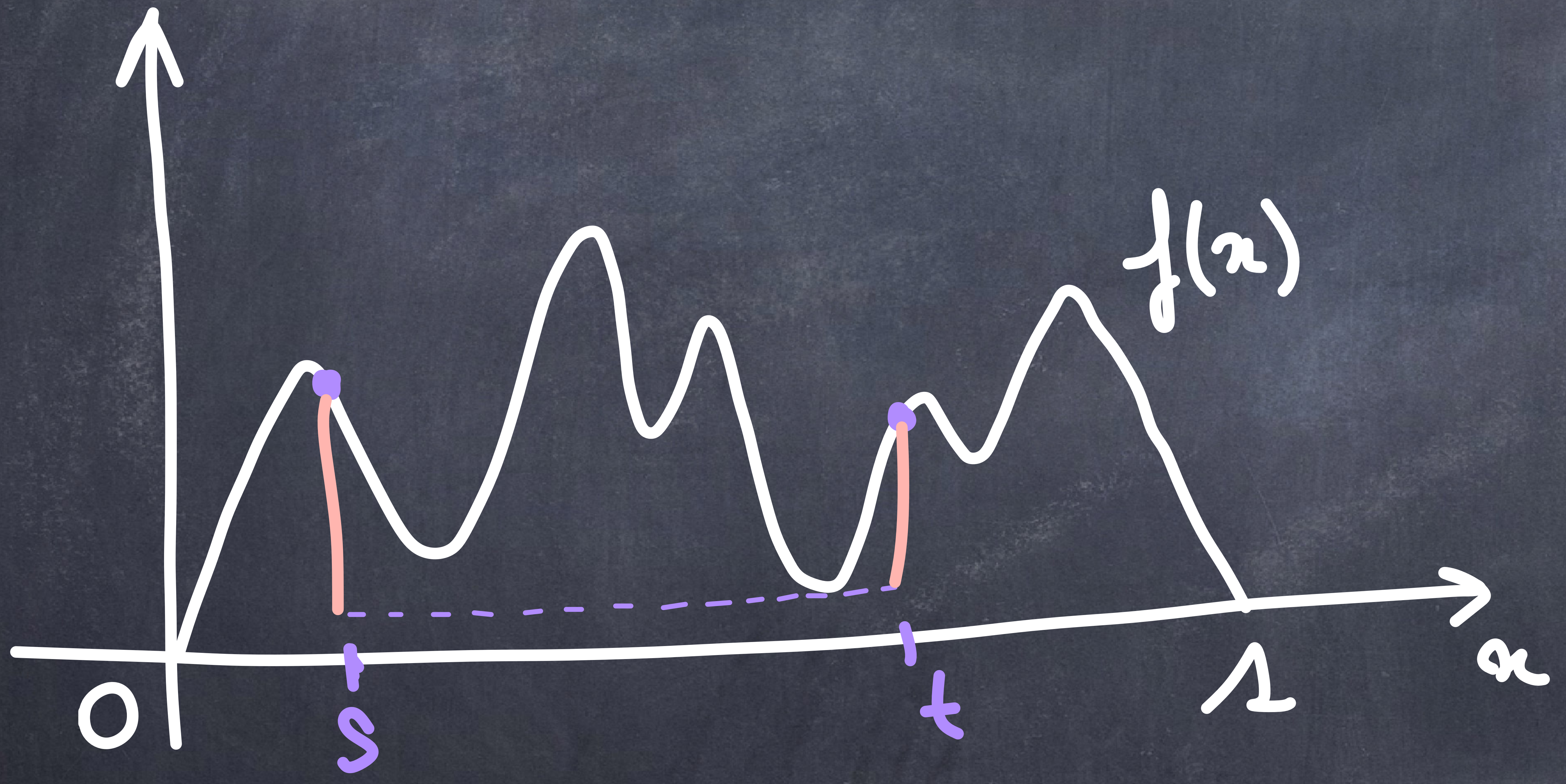


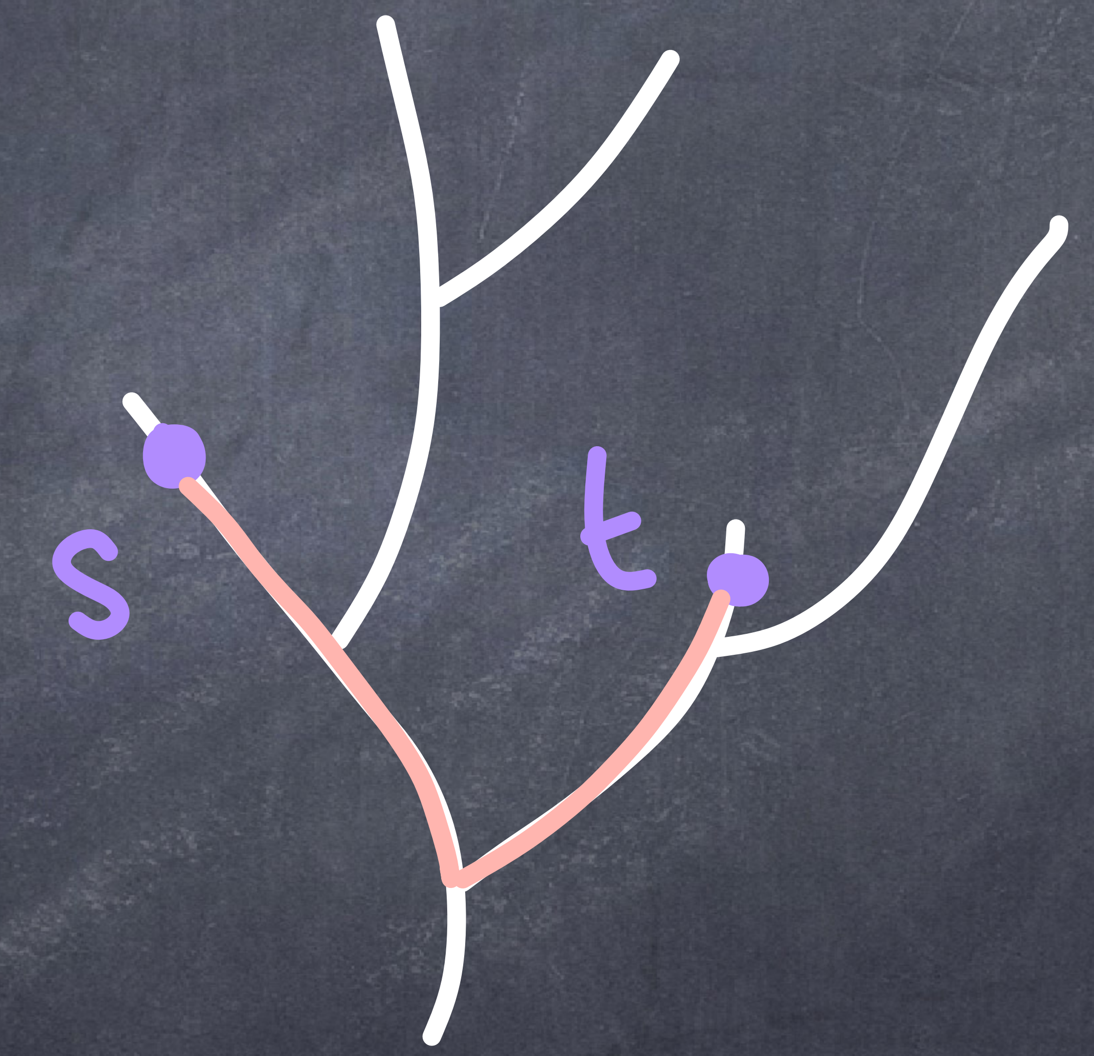
excursions in the upper half-plane $=$ trees

Trees as metric spaces.

A tree is a metric space: $\forall\, s,t \in [0,1]$

$$d(s,t) = f(s) + f(t) - 2 \min_{[s,t]} f$$





Brownian excursion = CRT

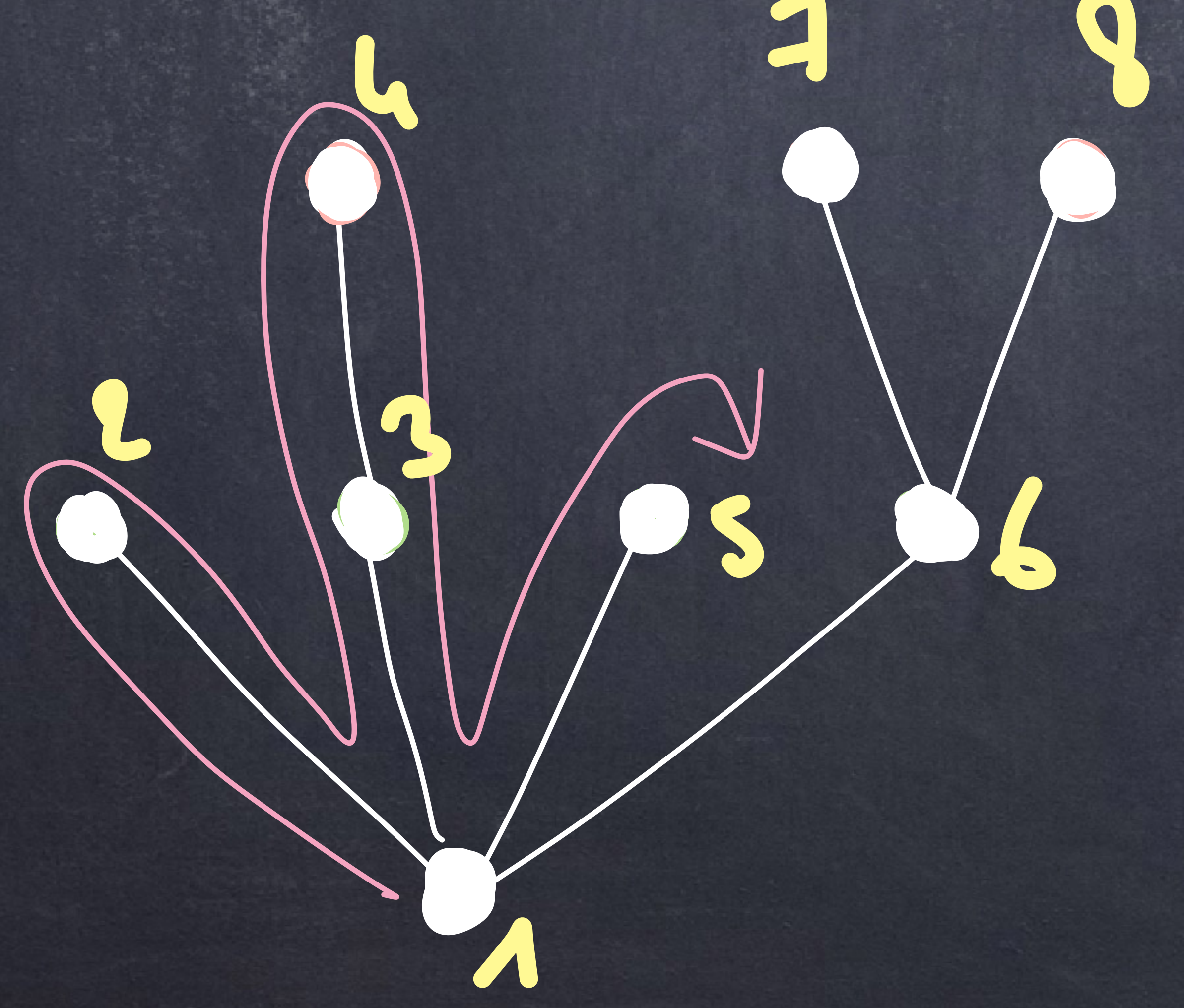
The Lukasiewicz path

Another excursion that characterises the tree:

$$L(n) = \sum_{i=1}^{n} (X_i - 1)$$

# of children of the ith node in the tree in "depth first order"

DFO = same as "contour" but do not repeat the nodes.



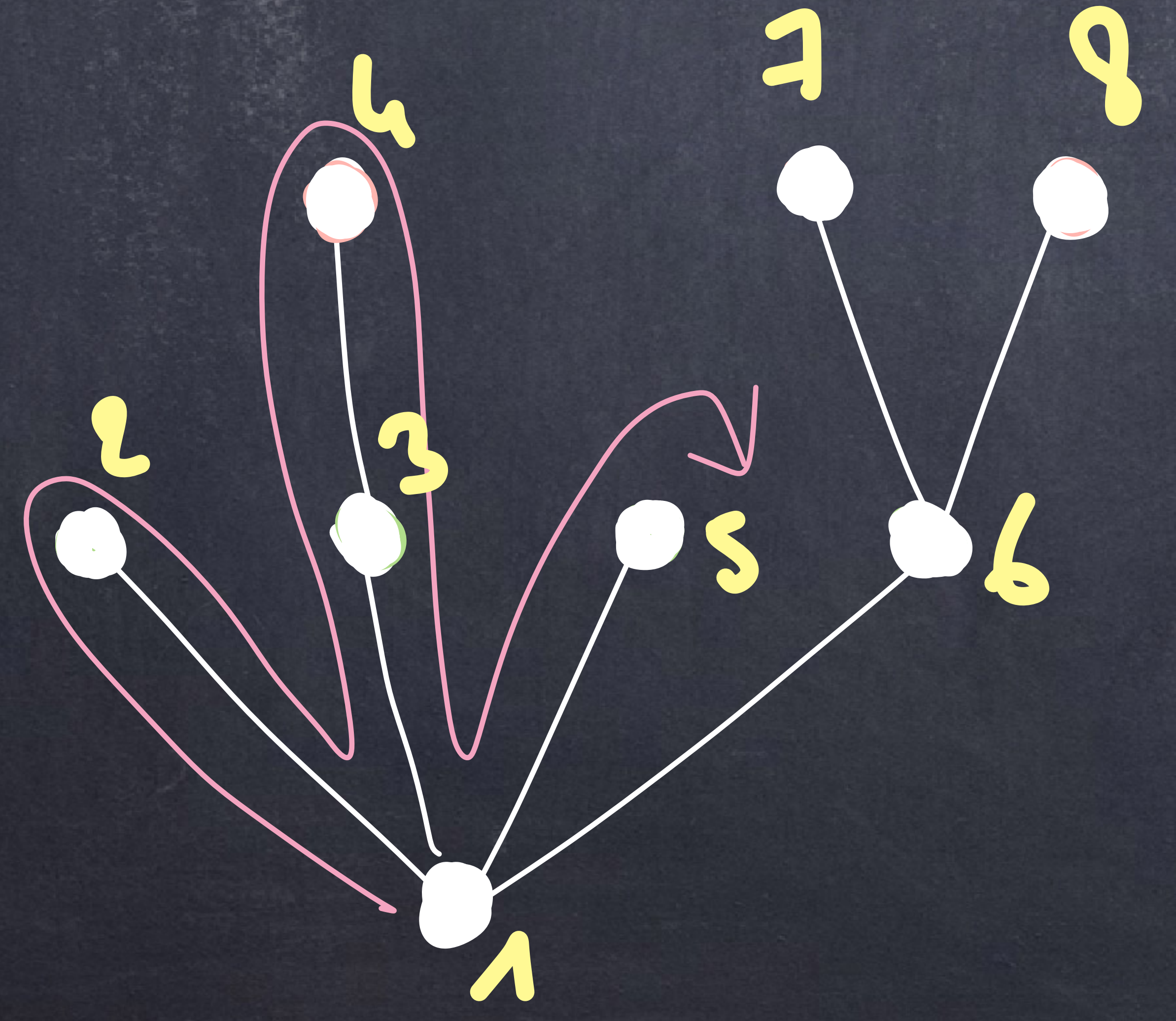
The Lukasiewicz path

Another excursion that characterises the tree:

$$L(n) = \sum_{i=1}^{n} (X_i - 1)$$

# of children of the i^{th} node in the tree in "depth first order"

DFO = same as "contour" but do not repeat the nodes.



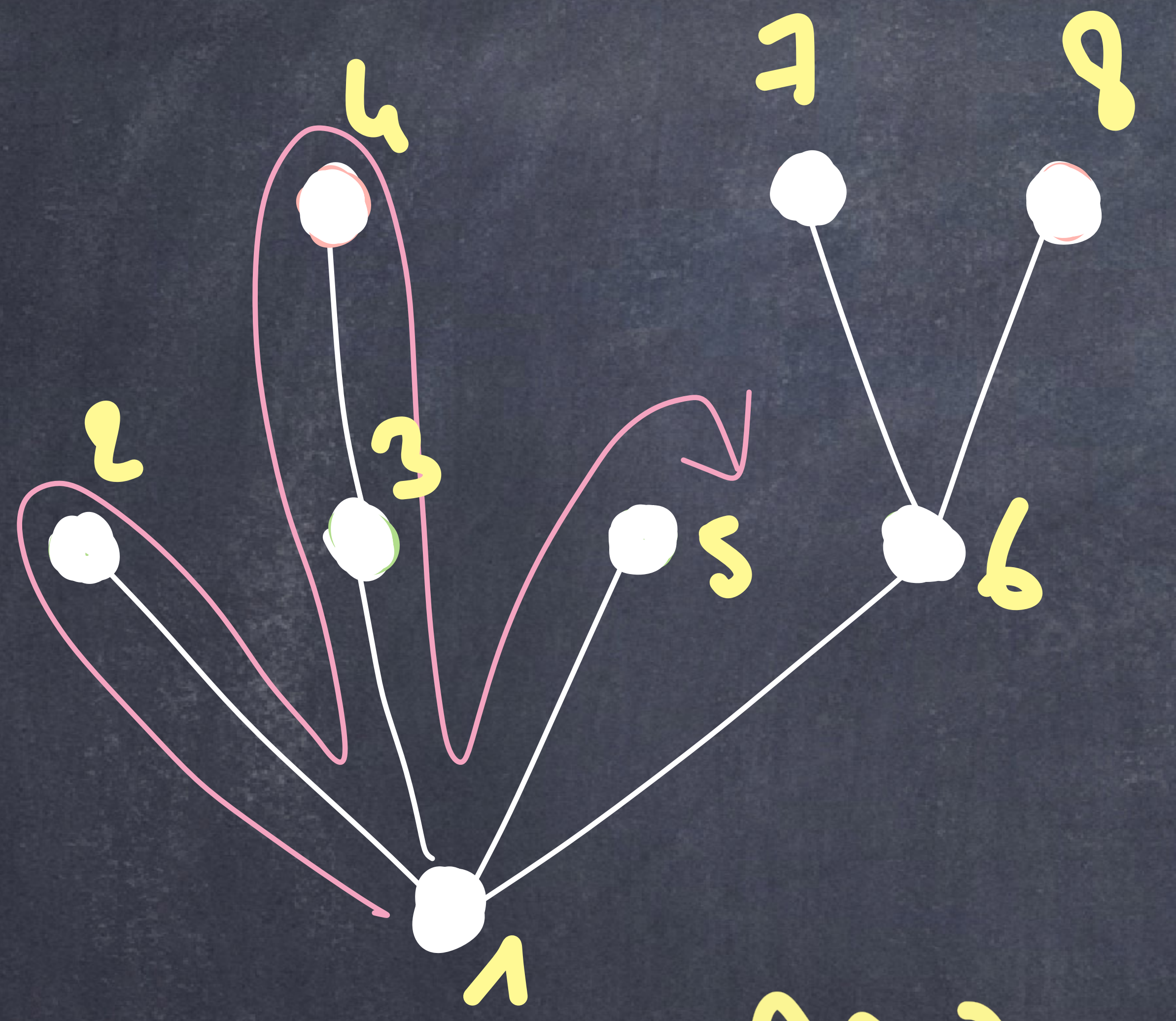
n-node tree satisfies:

- $L(n) = \left(\sum_{i=1}^{n} X_i\right) - n$
  $= (n-1) - n = -1$
- $L(k) \geq 0 \quad \forall\, 0 \leq k \leq n-1$

The Lukasiewicz path

$$L(k) = \sum_{i=0}^{k} (X_i - 1)$$

\# of children of the ith node in the tree in "depth first order"



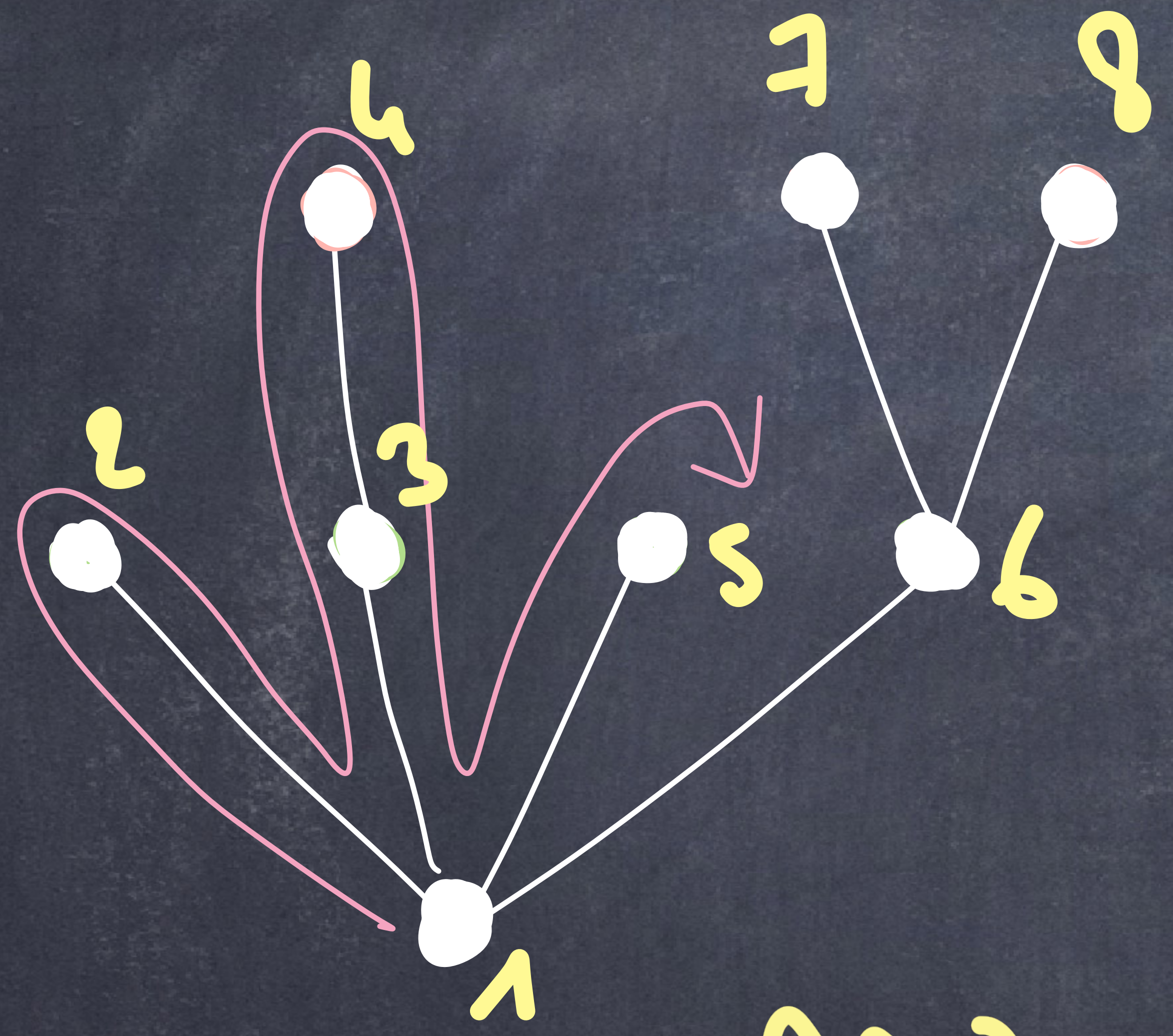
for a GW tree of offspring distribution μ, $(X_i)_{i \geq 1}$ are iid RVs of distribution μ.

$\rightsquigarrow$ $L(n)$ is a RW of iid increments of mean 0.

The Lukasiewicz path

$$L(k) = \sum_{i=0}^{k} (X_i - 1)$$

X_i: # of children of the ith node in the tree in "depth first order"



for a GW tree of offspring distribution μ, $(X_i)_{i \geq 1}$ are iid RVs of distribution μ.

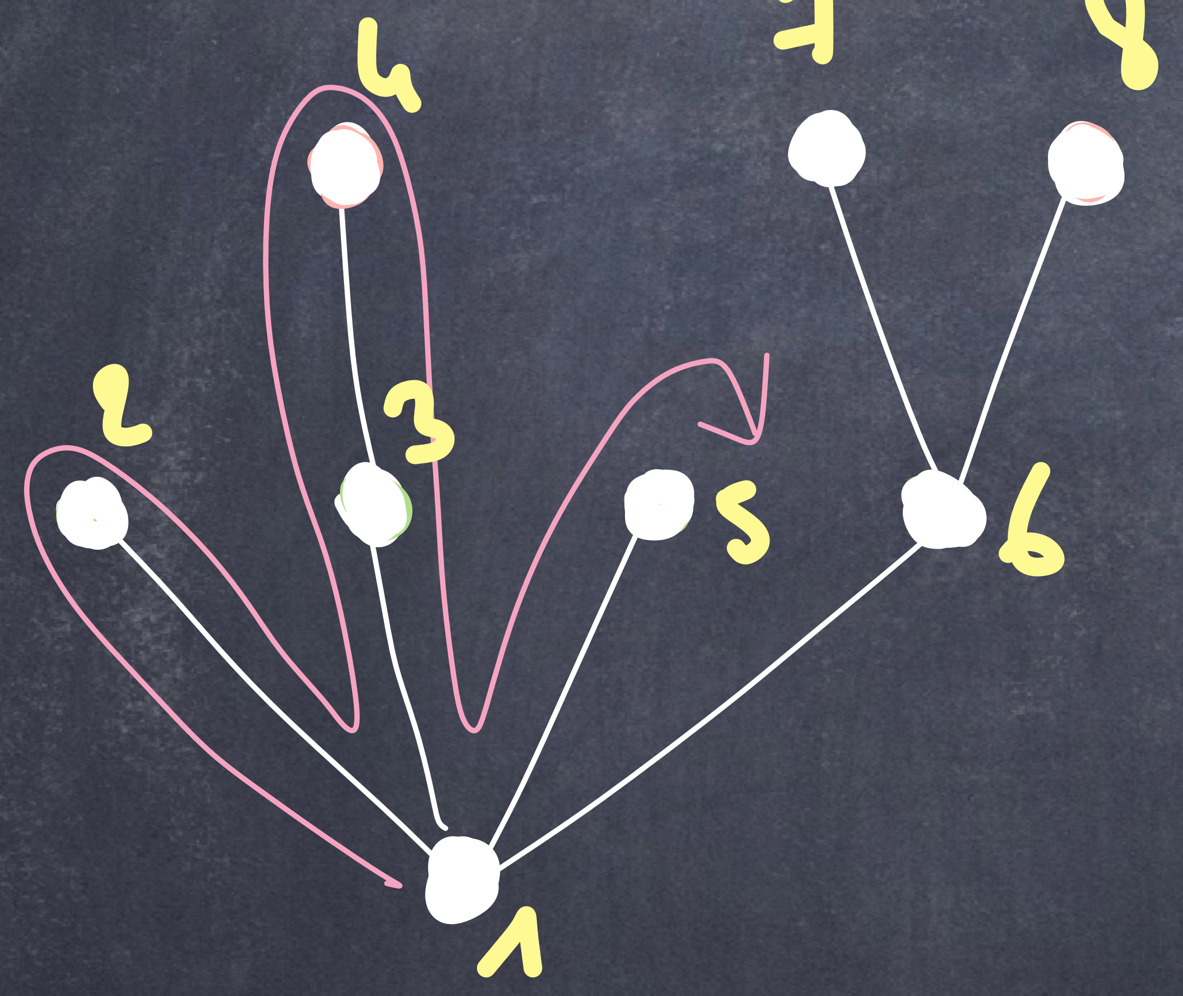
$\rightsquigarrow$ $L(n)$ is a RW of iid increments of mean 0.

$$\left(\frac{L(\lfloor tn \rfloor)}{\sigma \sqrt{n}} \right)_{t \in [0,1]} \xrightarrow[(d)]{n \uparrow \infty} BM \quad \text{(Skorokhod's topology)}$$

[Donsker's theorem]

The Łukasiewicz path

$$L(k) = \sum_{i=0}^{k} (X_i - 1)$$



Recall that, for an n-node tree,

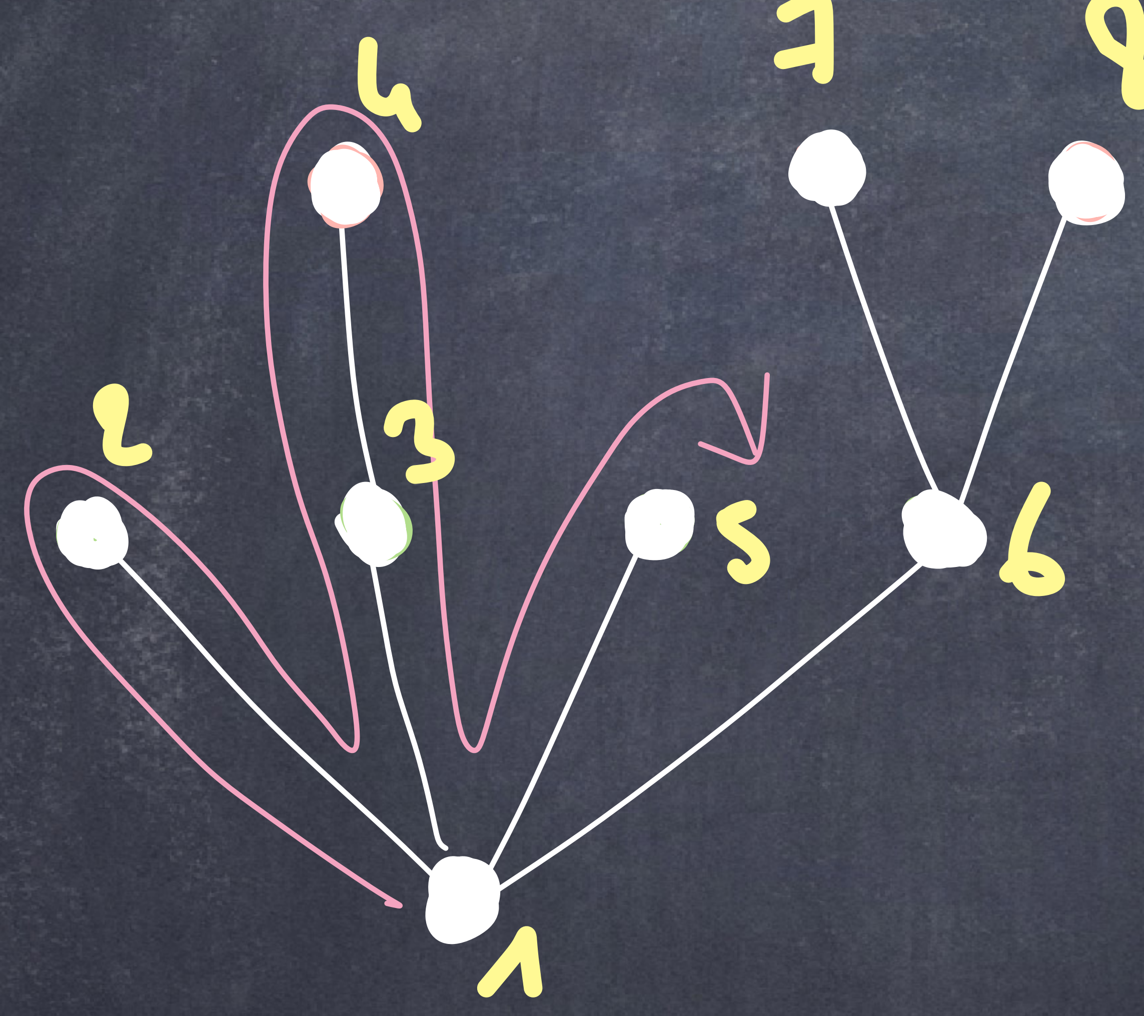
* $L(k) \geq 0 \quad \forall\, 0 \leq k \leq n-1$
* $L(n) = -1$

$$\left(\frac{L(\lfloor tn \rfloor)}{\sigma\sqrt{n}} \right)_{t\in[0,1]} \xrightarrow[(d)]{n\uparrow\infty} BM \quad \text{(Skorokhod's topology)}$$

[Donsker's theorem]

The Łukasiewicz path

$$L(k) = \sum_{i=0}^{k} (X_i - 1)$$



Recall that, for an n-node tree,

* $L(k) \geqslant 0 \quad \forall\, 0 \leq k \leq n-1$
* $L(n) = -1$

???

for large n, $L(\lfloor nt \rfloor)$ could be < 0 !

BM typically can be < 0 ⚠

$$\left(\frac{L(\lfloor tn \rfloor)}{\sigma\sqrt{n}}\right)_{t\in[0,1]} \xrightarrow[(d)]{n\uparrow\infty} BM \quad \text{(Skorokhod's topology)}$$

[Donsker's theorem]

The Lukasiewicz path

Idea: consider an infinite forest of iid $GW(\mu)$.

$-\min_{i\le n} L(i) = \#$ previous trees in the forest

$$\left(\frac{L(\lfloor tn\rfloor)}{\sigma\sqrt{n}}\right)_{t\in[0,1]} \xrightarrow[(d)]{n\to\infty} BM$$ (Skorokhod's topology)

[Donsker's theorem]

The Lukasiewicz path

Idea: consider an infinite forest of iid $GW(\mu)$.

Reflect the walk above the running minimum.

$$\Pi(n) = \min_{0\leq i\leq n} L(i) \quad \forall n\geq 0.$$

$$\left(\frac{L(\lfloor tn\rfloor) - \Pi(\lfloor tn\rfloor)}{\sigma\sqrt{n}}\right)_{t\in[0,1]} \xrightarrow[(d)]{n\uparrow\infty} |B\Pi|$$

(Skorokhod's topology)

The Łukasiewicz path

Idea: consider an infinite forest of iid $GW(\mu)$.

But... convergence of the Łukasiewicz path is not enough to prove cv in the GH sense! :(

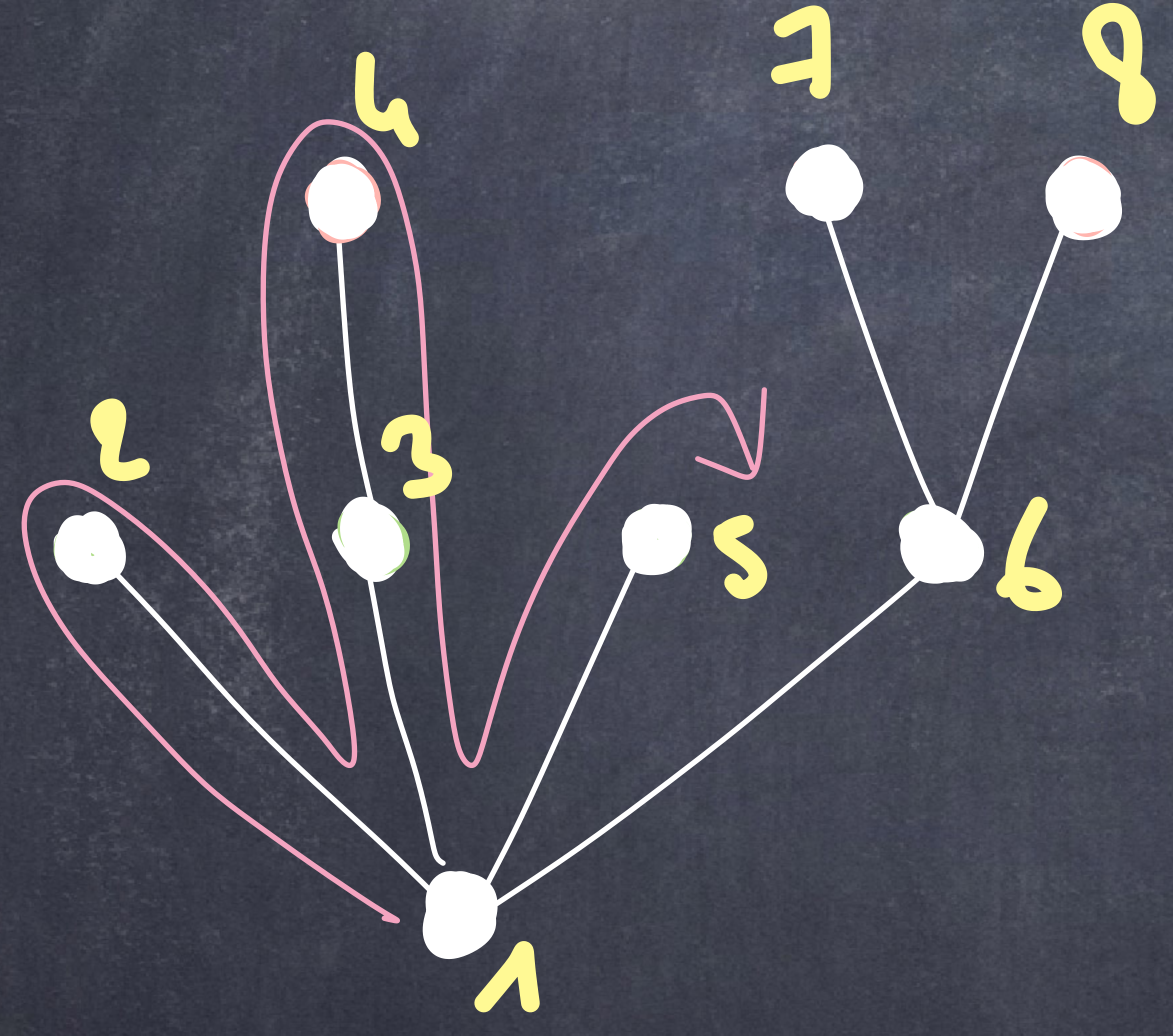
$$\left(\frac{L(\lfloor tn \rfloor) - \Pi(\lfloor tn \rfloor)}{\sigma\sqrt{n}} \right)_{t\in[0,1]} \xrightarrow[(d)]{n\uparrow\infty} |B\Pi|$$

(Skorokhod's topology)

The height process

Yet another function that characterises a tree / forest:

$H(n) =$ height of the n^{th} node in DFO

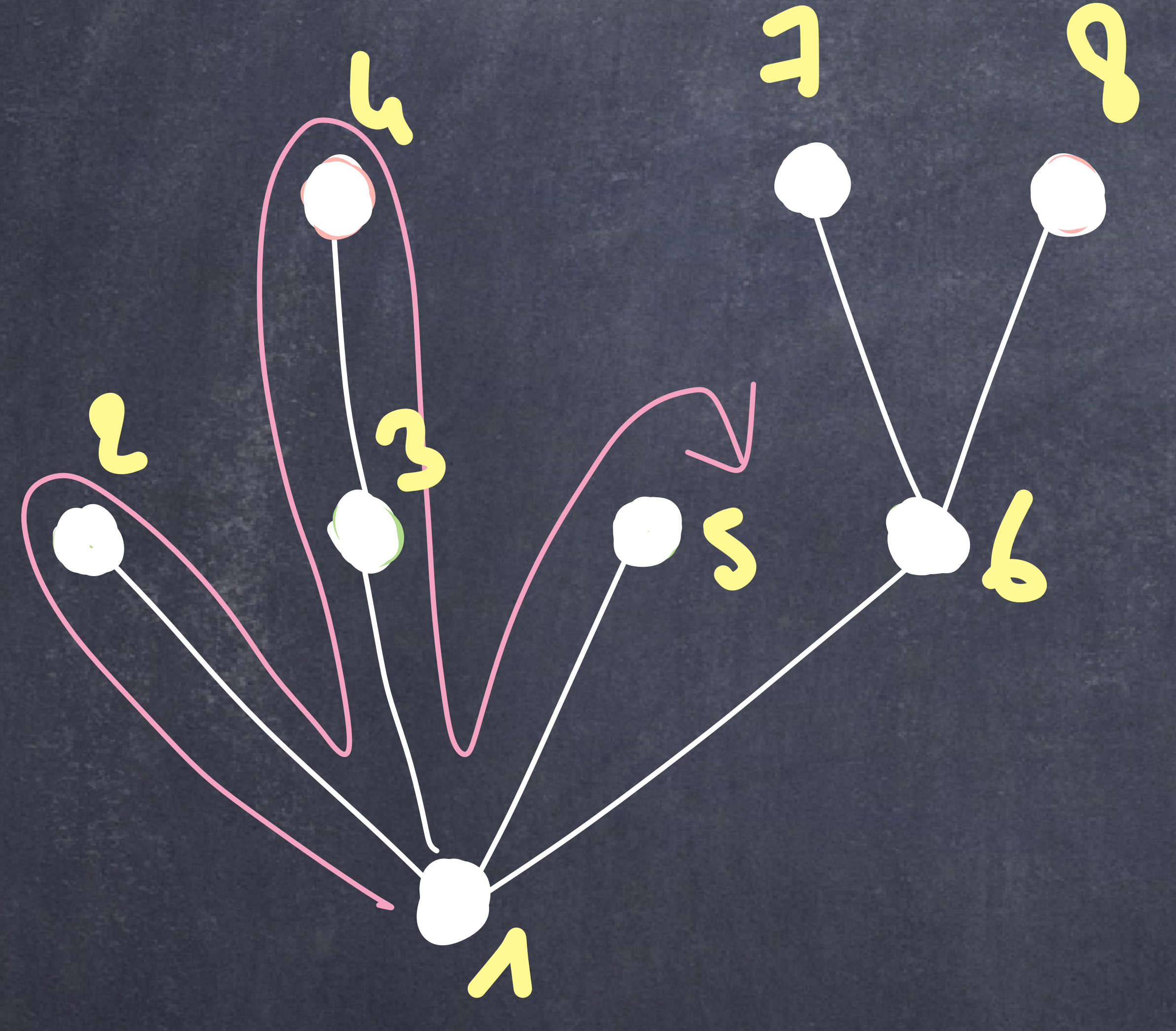


$=$

The height process

Yet another function that characterises a tree / forest:

$H(n)$ = height of the n^{th} node in DFO



=

⚠ if $\left(\dfrac{H(\lfloor nt \rfloor)}{\alpha\sqrt{n}}\right)_{t\in[0,1]} \xrightarrow[n\uparrow\infty]{(d)} |BM|$ then $\mathcal{T}_n \longrightarrow$ CRT !

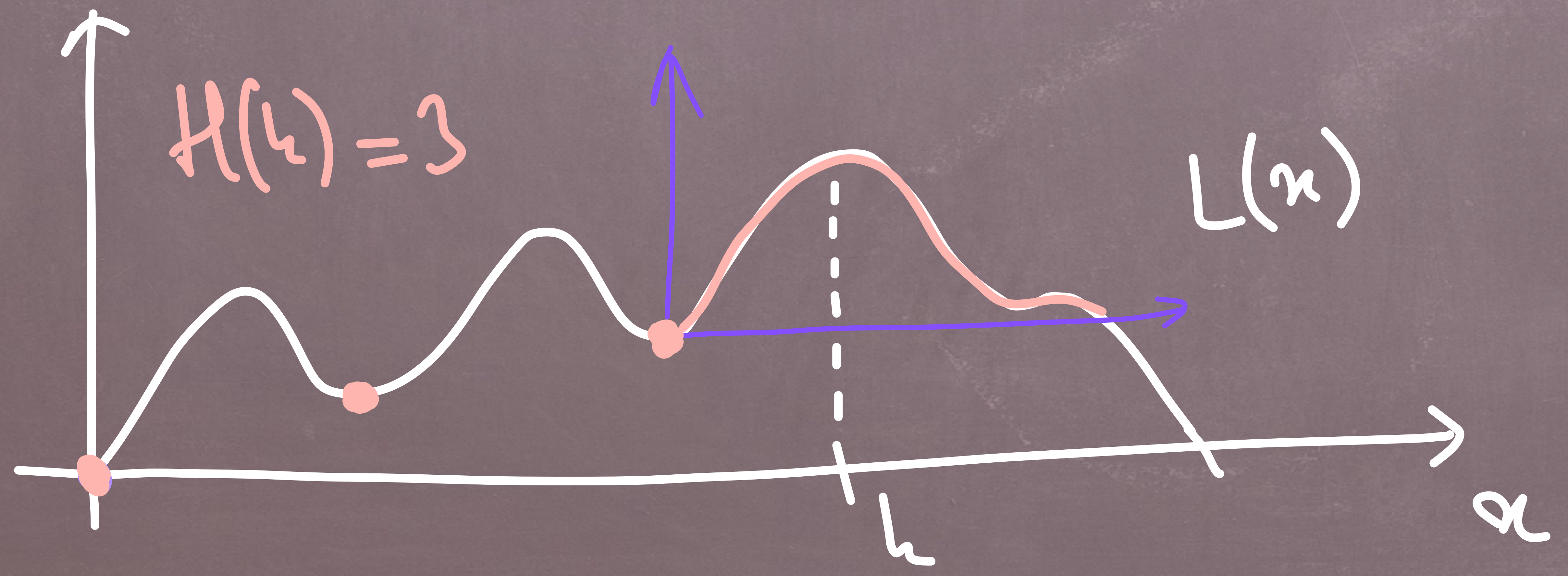
The height process

Yet another function that characterises a tree / forest :

$H(n) =$ height of the n^{th} node in DFO

Lemma : for all $n \geq 1$, $0 \leq k \leq n-1$,

$$H(k) = \# \left\{ 0 \leq i \leq k-1 \; : \; L(i) = \min_{i \leq j \leq k} L(j) \right\}$$



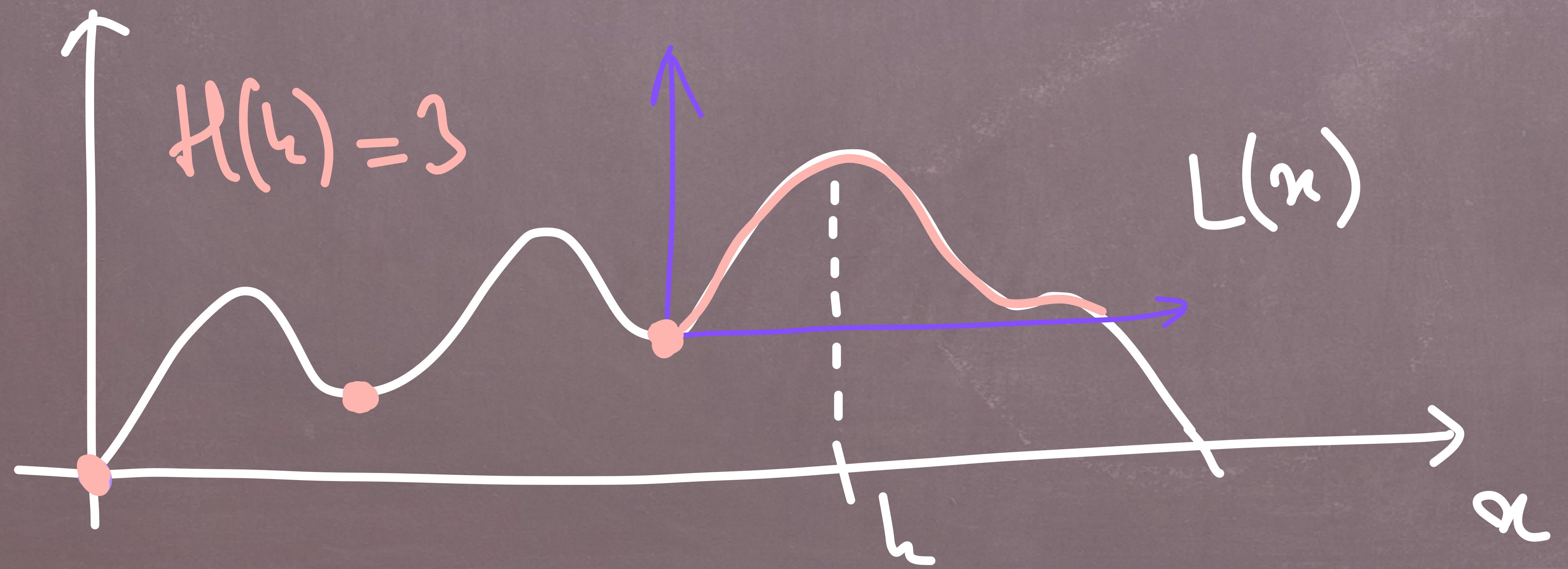
The height process

Yet another function that characterises a tree / forest:

$H(n) =$ height of the n^{th} node in DFO

lemma: for all $n \geq 1$, $0 \leq k \leq n-1$,

$$H(k) = \#\{0 \leq i \leq k-1 : L(i) = \min_{i \leq j \leq k} L(j)\}$$



⇝ If $L \longrightarrow |B\Pi|$

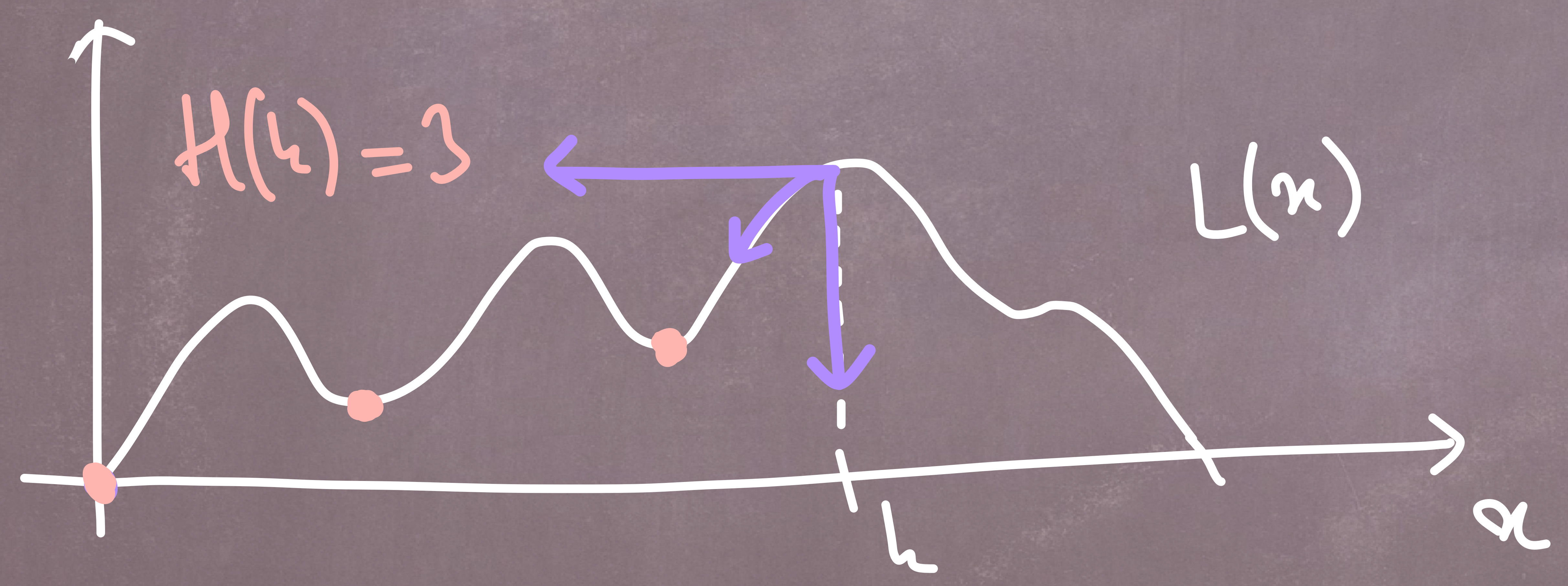
then $H \longrightarrow |B\Pi|$ too!

:)

The height process

Yet another function that characterises a tree / forest:

$H(n) =$ height of the n^{th} node in DFO



$\rightsquigarrow$ If $L \to |BM|$
then $H \to |BM|$ too !

⚠ The proof relies on the fact that the walk that takes the same steps as L in reverse order from time $k \overset{(d)}{=} L$

Proof of convergence of $GW(\mu)$ to CRT

Th (Aldous '91): if $\sigma^2 := Var(\mu) < \infty$, then

$$\frac{1}{\sigma\sqrt{n}} \mathcal{T}^{(n)} \xrightarrow[n\uparrow\infty]{(d)} \text{CRT} \quad \text{for the GH metric.}$$

1. The Łukasiewicz path cv to BM [Donsker]
2. The height process cv to BM
3. The largest excursion = a tree conditioned to be large.

Galton-Watson trees in random environment

$(\mu_n)_{n\geq 0}$ = a sequence of iid probability distributions on $\mathbb{N}$

$\rightsquigarrow$ we use μ_n for the offspring distribution of the individuals of generation n.

Galton-Watson trees in random environment

$(\mu_n)_{n\geq 0}$ = a sequence of iid probability distributions on $\mathbb{N}$

$\leadsto$ we use μ_n for the offspring distribution of the individuals of generation n.

We assume that, a.s., $\mu_0(0) < 1$ and $\bar{\mu}_0 = \sum_{k\geq 0} k\mu_0(k) \in (0,\infty)$

Galton–Watson trees in random environment

$(\mu_n)_{n\geq 0}$ = a sequence of iid probability distributions on $\mathbb{N}$

⇝ we use μ_n for the offspring distribution of the individuals of generation n.

We assume that, a.s., $\mu_0(0) < 1$ and $\bar{\mu}_0 = \sum_{k\geq 0} k\mu_0(k) \in (0,\infty)$

Th (Kersting &) Assume that $\log \bar{\mu}_0$ is integrable.

Then the GWRE is super-critical, critical or subcritical according to whether $\mathbb{E}[\log \bar{\mu}_0] > 0$, $\mathbb{E}[\log \bar{\mu}_0] = 0$ or $\mathbb{E}[\log \bar{\mu}_0] < 0$.

GWREs : extinction probability

$\underline{\underline{\text{NB}}}$: $\mathbb{E}[\log \bar{\mu}_o] \leq \log \mathbb{E}[\bar{\mu}_o]$ [Jensen's inequality]

Thus a GWRE can be sub-critical & still satisfy $\mathbb{E}\bar{\mu}_o > 1$, which implies $\mathbb{E} Z_n \longrightarrow \infty$.

GWREs : extinction probability

NB : $\mathbb{E}[\log \bar{\mu}_0] \leq \log \mathbb{E}[\bar{\mu}_0]$ [Jensen's inequality]

Thus a GWRE can be sub-critical & still satisfy $\mathbb{E}\bar{\mu}_0 > 1$, which implies $\mathbb{E} Z_n \longrightarrow \infty$.

Example:

$$\mu_0 = \begin{cases} (1-\varepsilon)\delta_0 + \varepsilon\delta_1 & \text{with probability } 1/2 \\ \delta_2 & \text{with probability } 1/2 \end{cases}$$

GWREs : extinction probability

NB : $\mathbb{E}[\log \bar{\mu}_0] \leq \log \mathbb{E}[\bar{\mu}_0]$ [Jensen's inequality]

Thus a GWRE can be sub-critical & still satisfy $\mathbb{E}\bar{\mu}_0 > 1$, which implies $\mathbb{E} Z_n \longrightarrow \infty$.

Example: $\mu_0 = \begin{cases} (1-\varepsilon)\delta_0 + \varepsilon\delta_1 & \text{with probability } 1/2 \\ \delta_2 & \text{with probability } 1/2 \end{cases}$

Thus, $\bar{\mu}_0 = \begin{cases} \varepsilon & \text{w.p. } 1/2 \\ 2 & \text{w.p. } 1/2 \end{cases} \Longrightarrow \begin{cases} \mathbb{E}\bar{\mu}_0 = 1 + \frac{\varepsilon}{2} > 1 \\ \mathbb{E}\log\bar{\mu}_0 = \frac{1}{2}\log\varepsilon + \frac{1}{2}\log 2 < 0 \text{ if } \varepsilon < \frac{1}{2} \end{cases}$

GWREs : extinction probability

Th (Kersting 20) Assume that $\log \bar{\mu}_0$ is integrable.
Then the GWRE is super-critical, critical or subcritical according to whether $\mathbb{E}[\log \bar{\mu}_0] > 0$, $\mathbb{E}[\log \bar{\mu}_0] = 0$ or $\mathbb{E}[\log \bar{\mu}_0] < 0$.

sub-critical & critical $\rightsquigarrow$ $\boxed{\text{a.s. extinction}}$

super-critical + $\mathbb{E}[|\log(1-\mu_0(0))|] < \infty$ $\rightsquigarrow$ $\boxed{\text{possible survival}}$

We say that the process is strictly critical if $\log \bar{\mu}_0 = 0$ a.s.
i.e. $\bar{\mu}_0 = 1$ a.s.

GWREs : recent results in the literature

"strictly critical GWRE behave as critical GW trees"

* survival probability & population size (Kersting 18)
* population size (Cardona-Tobón & Palau 21)
* reduced process (Kersting 22)
* genealogy of a k-sample (Harris, Palau & Pardo 22)
  (Boenkost, Foutel-Rodier, Schertzer 22)

GWREs : scaling limit

Th (Bonchon–Kejan, Kious, n 21) Assume that $\bar{\mu}_0 = 0$ a.s.

Let $\sigma^2 := E\sigma_0^2 < \infty$. Let $\mathcal{T}_n$ be the GWRE$(\mu_n)_{n\geq 0}$ conditioned on having total population at least n. Then

$$\frac{1}{\sigma\sqrt{n}}\mathcal{T}_n \xrightarrow[(d)]{n\uparrow\infty} \text{CRT}$$

for the GH metric.

$$\sigma_0^2 = \sum_{k\geq 0}(k-\bar{\mu}_0)^2\mu_0(k)$$

GWREs : scaling limit

Th (Bonchon – Kuijer, Kious, ? 21) Assume that $\bar{\mu}_0 = 0$ a.s.

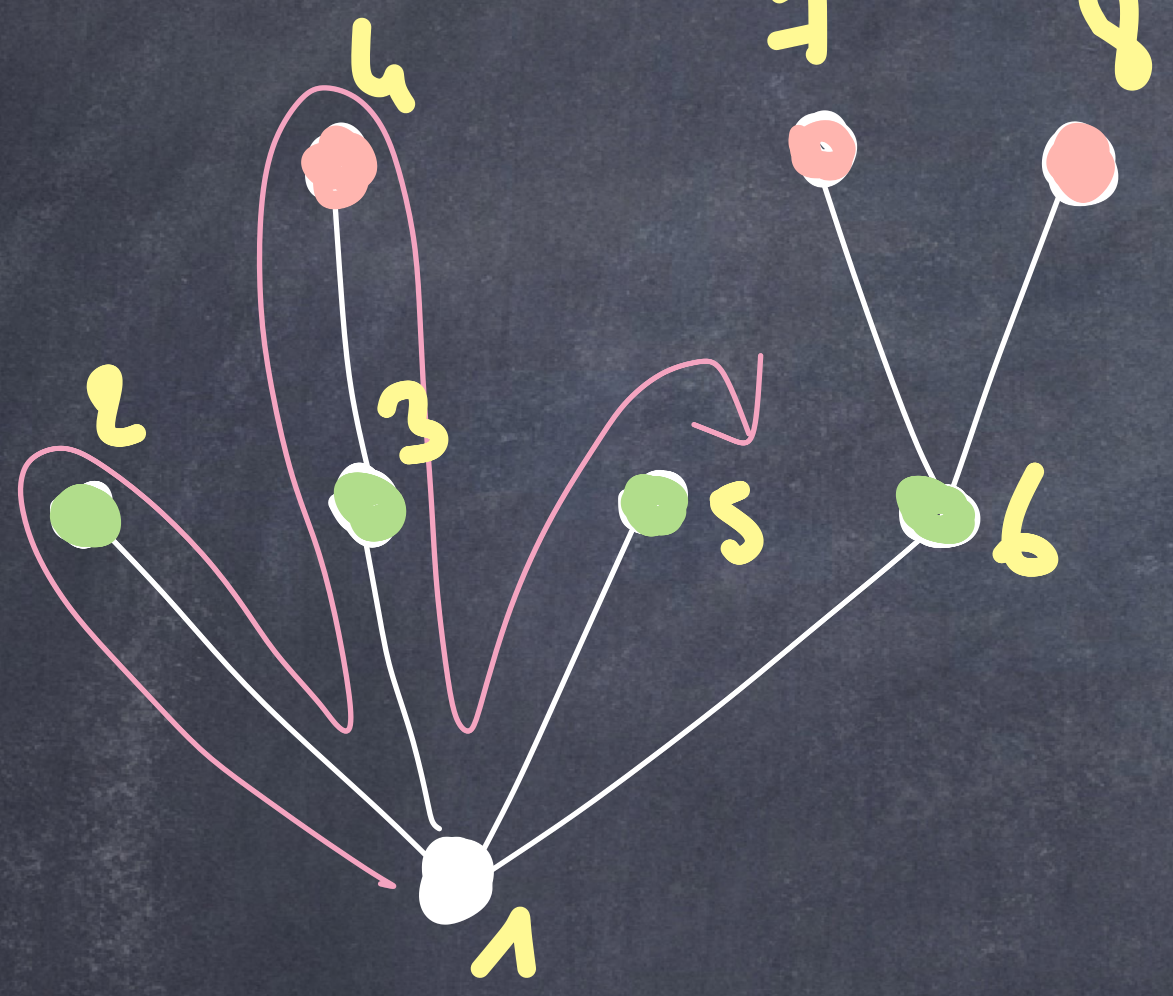
$$\boxed{\text{\& } \sigma^2 := E\sigma_0^2 < \infty. \text{ Let } \mathcal{T}_n \text{ be the GWRE}(\mu_n)_{n\geq 0} \text{ conditioned on having total population at least } n. \text{ Then } \frac{1}{\sigma\sqrt{n}}\mathcal{T}_n \xrightarrow[(d)]{n\uparrow\infty} \text{CRT for the GH metric.}}$$

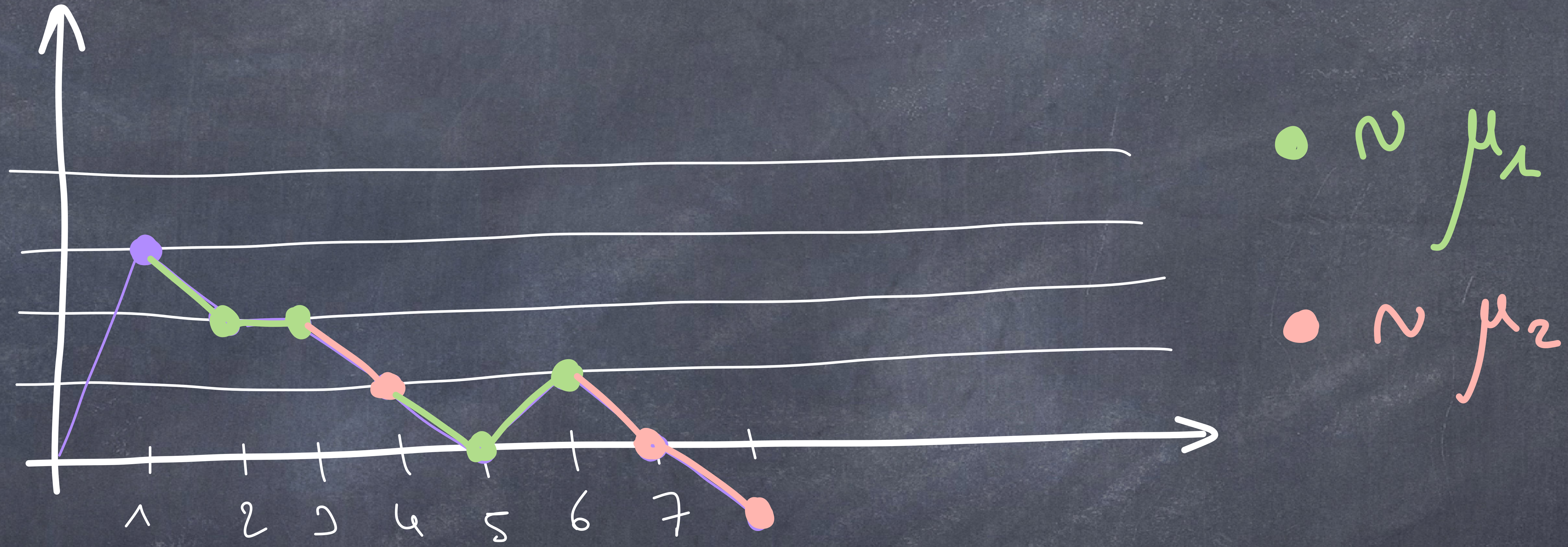
$\sigma_0^2 = \sum_{k\geq 0}(k-\bar{\mu}_0)^2 \mu_0(k)$

Same as the classical GW tree... but the proof is much more involved! Mainly because L is no longer a RW with iid increments.

GWREs : scaling limit

L is no longer a RW with iid increments...

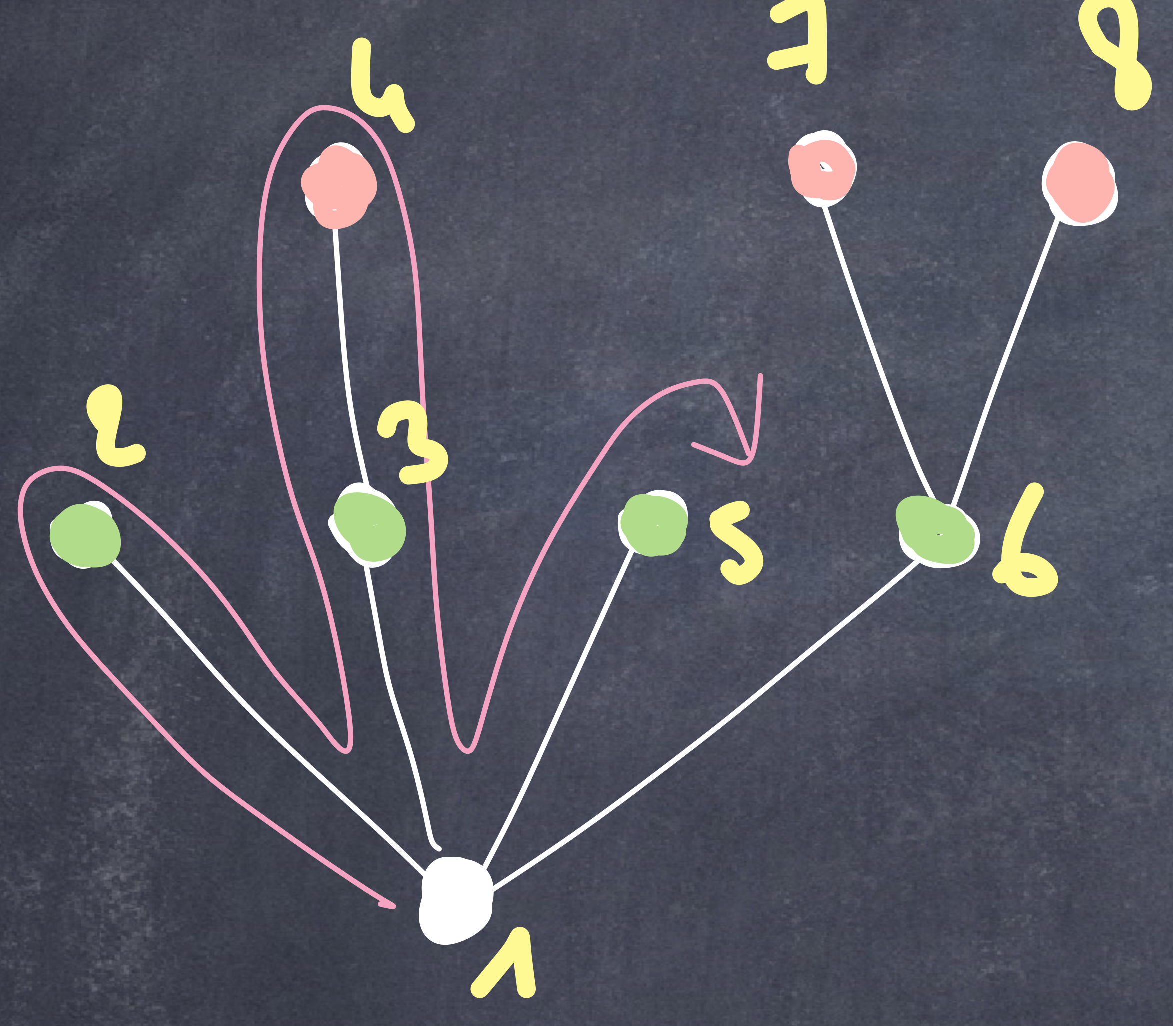


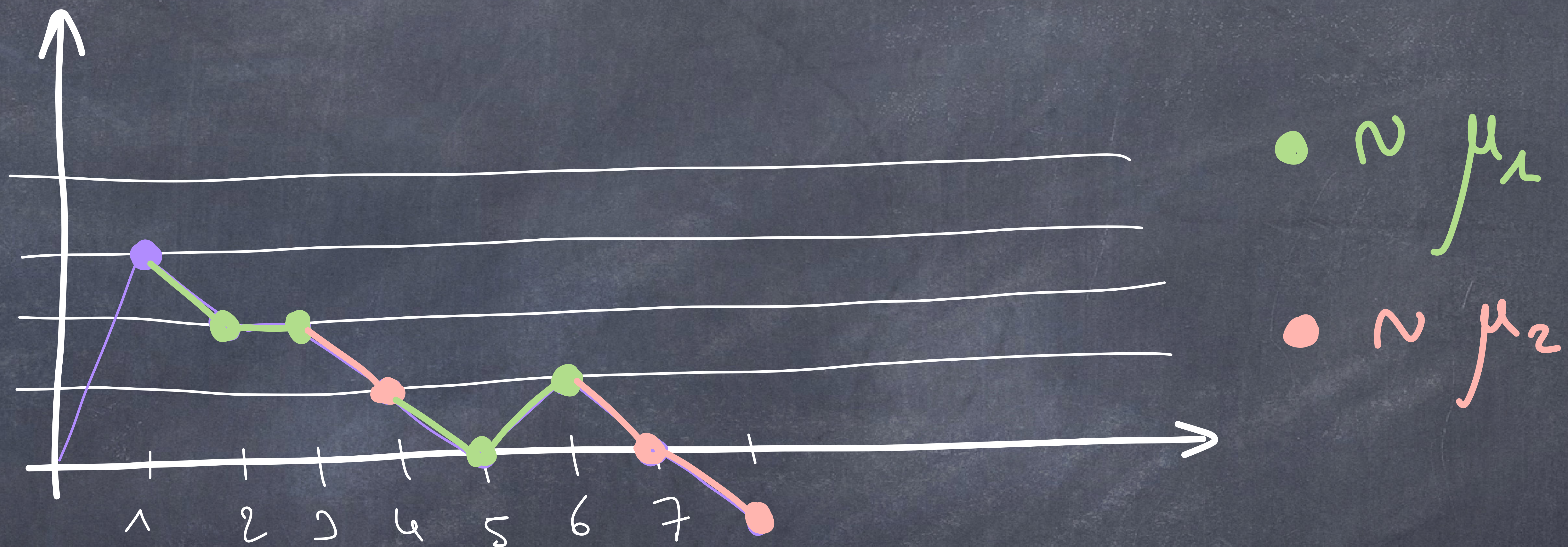


in fact, it is no longer Markovian!

GWREs : scaling limit

L is no longer a RW with iid increments...





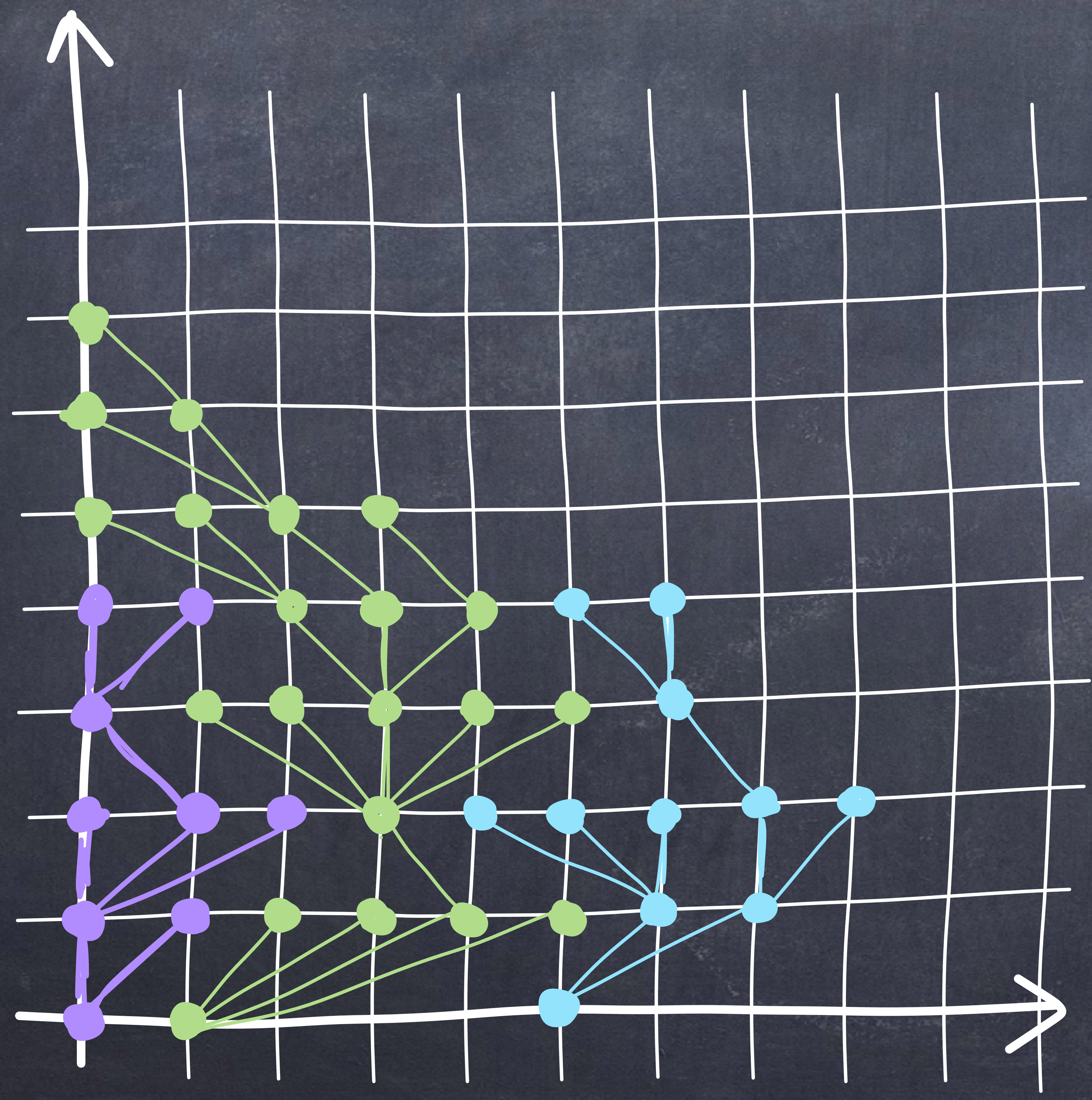
in fact, it is no longer Markovian!

And the walk taking steps in reverse order $\underset{(d)}{\neq}$ L !

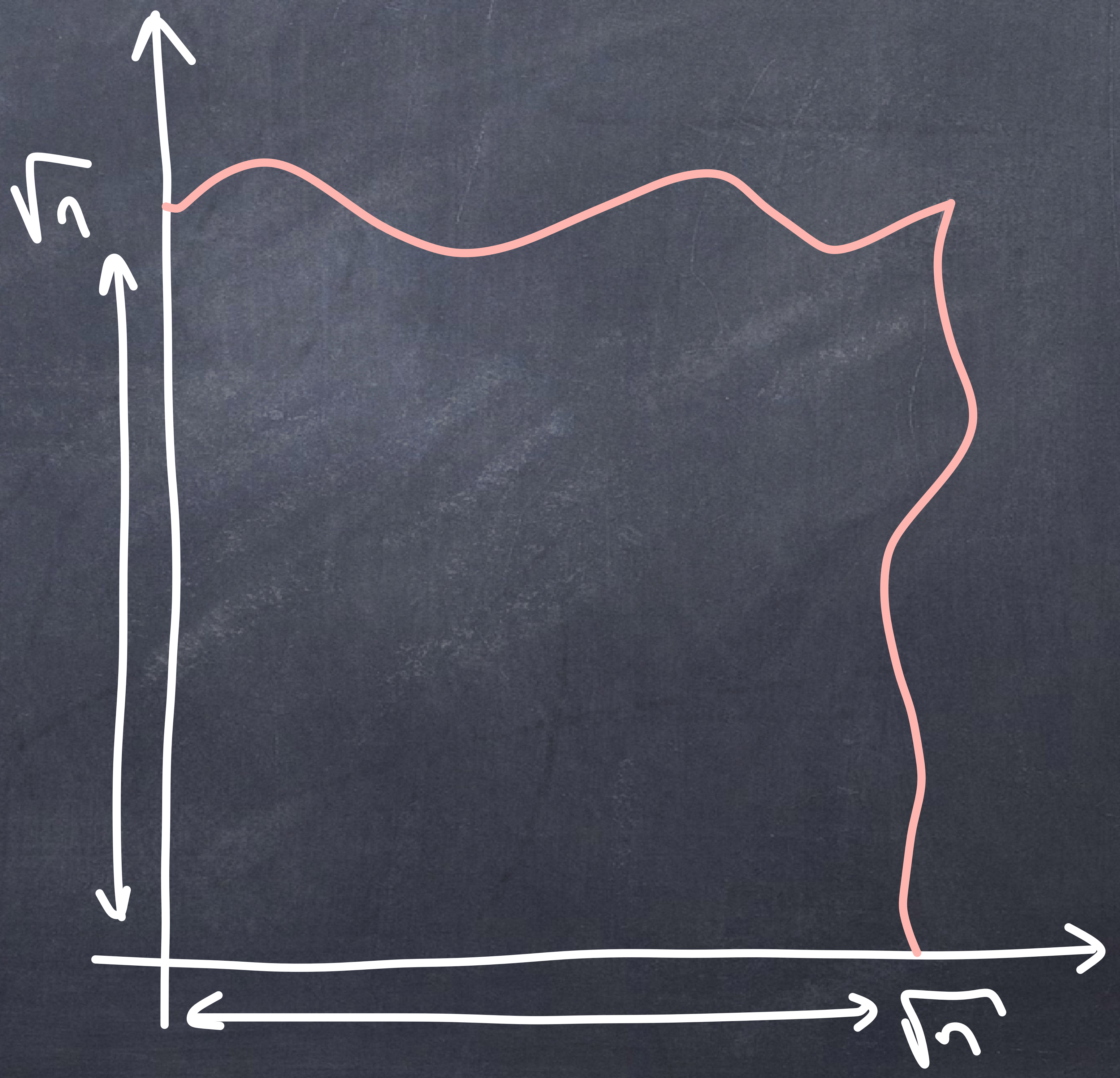
Ideas of the proof: ① Convergence of the Lukasiewicz path

$$L(n) = \sum_{i=1}^{n} (X_i - 1) \qquad (\forall n \geq 1)$$

children of the i^{th} node.



$\approx$ large n



Ideas of the proof: ① Convergence of the Łukasiewicz path

$$L(n) = \sum_{i=1}^{n} (X_i - 1) \qquad (\forall n \geq 1)$$

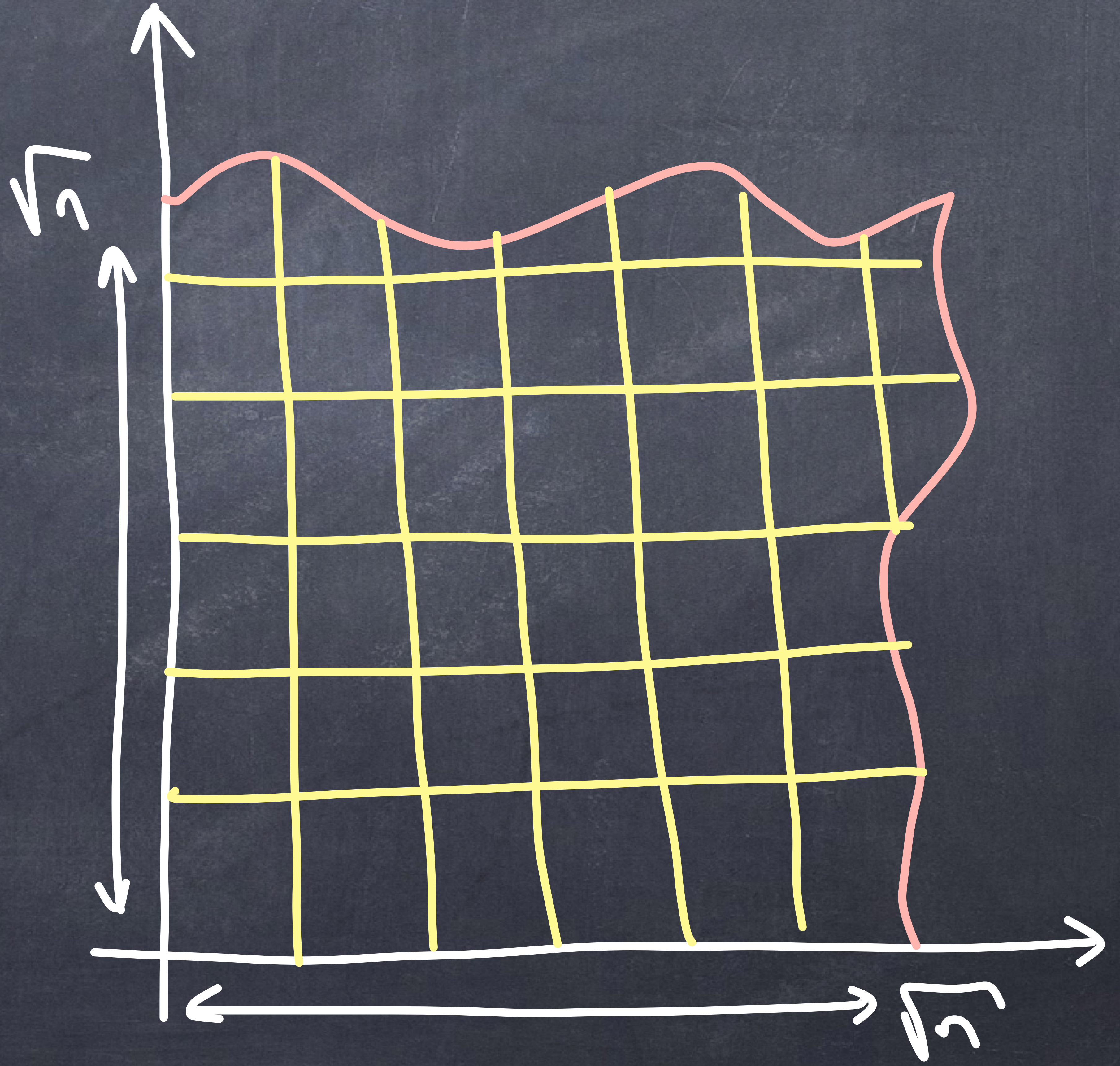
\# children of the i^{th} node.

We use a martingale CLT to prove

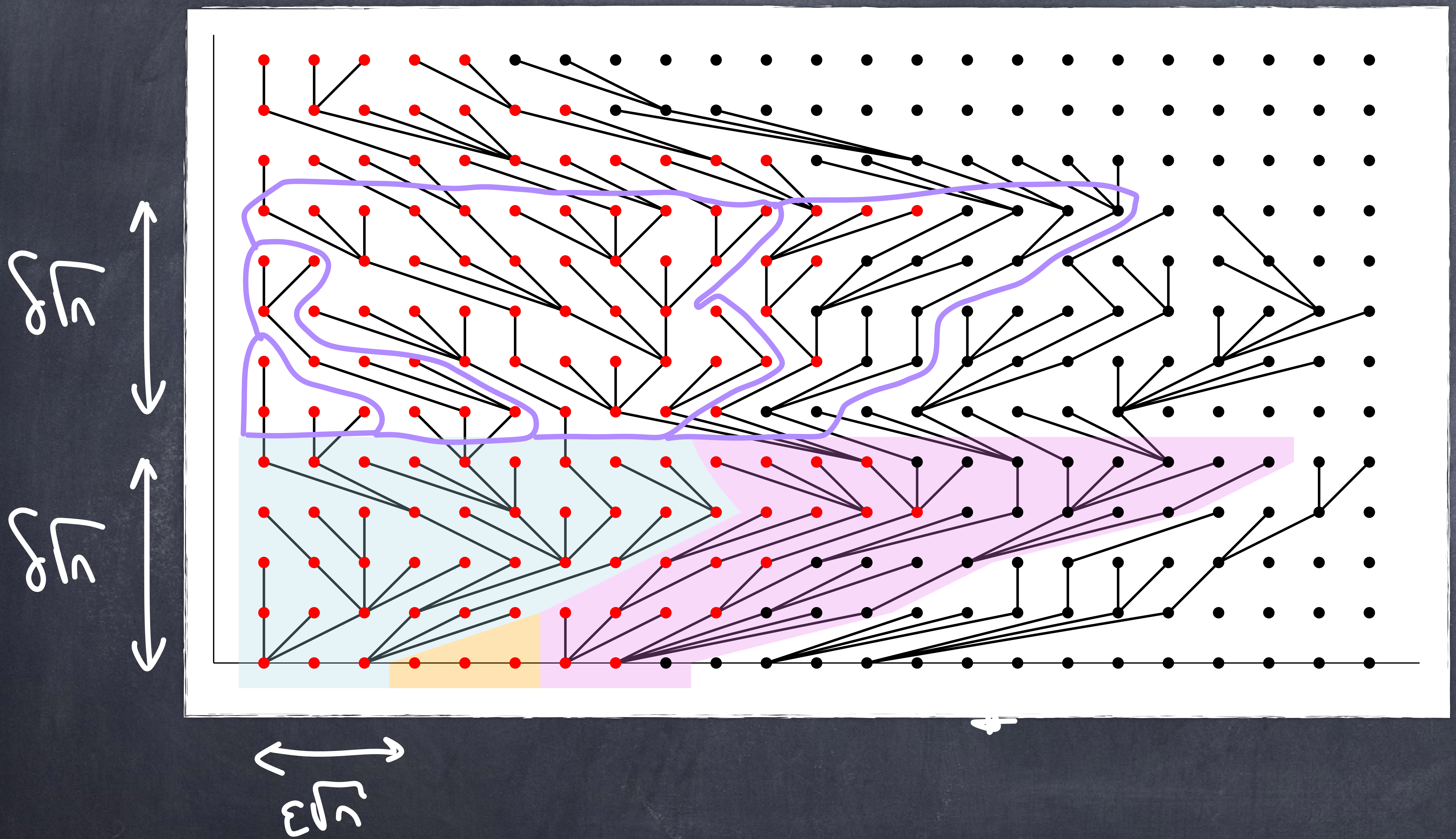
$$\left(\frac{L(\lfloor nt \rfloor)}{\sigma \sqrt{n}} \right)_{t \in [0,1]} \xrightarrow{(d)} B\Pi$$

Main condition :

$$\frac{1}{n} \sum_{i=1}^{n} \sigma^2_{H(i)} \longrightarrow \sigma^2$$

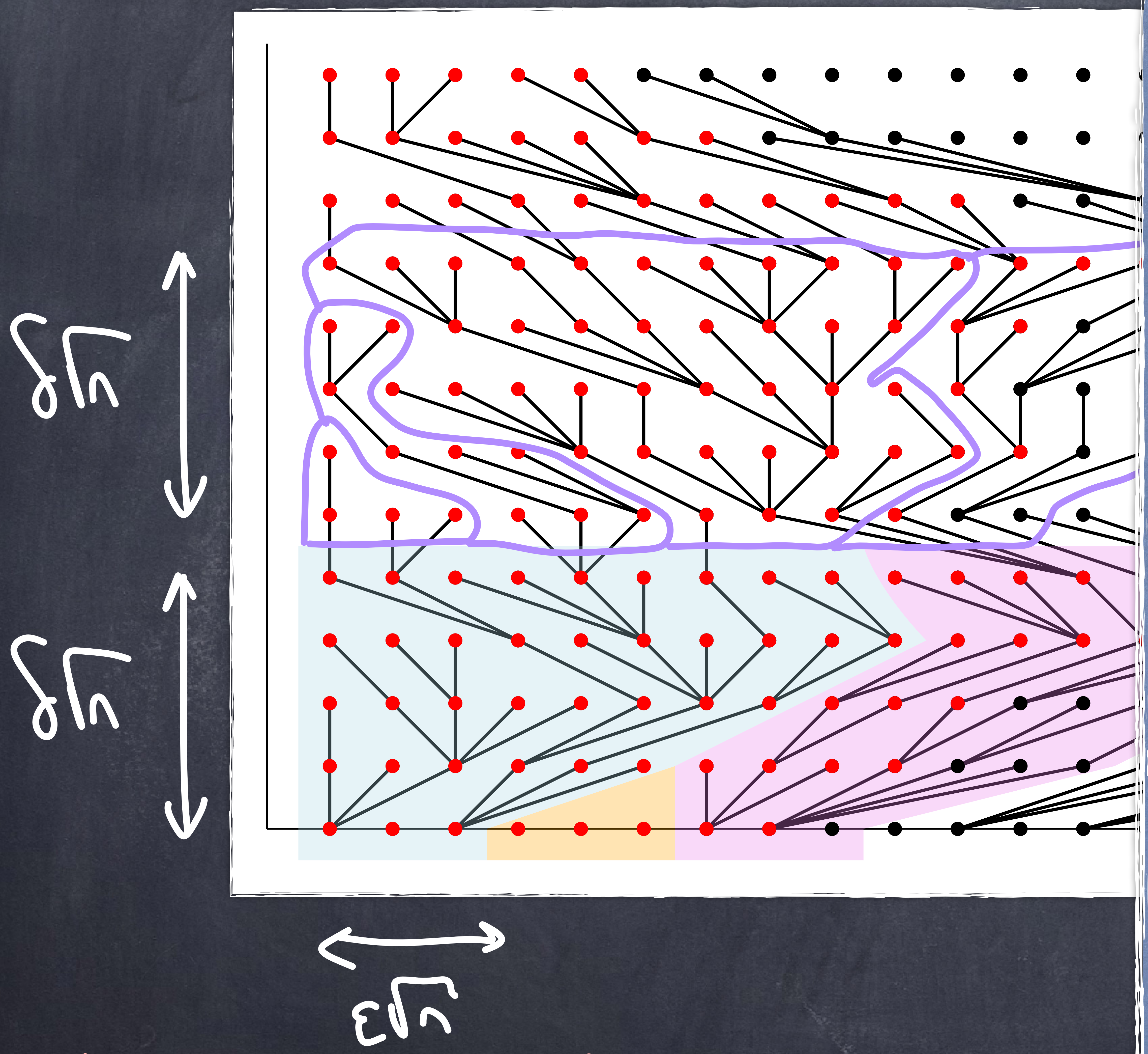


Ideas of the proof: 1 Convergence of the Łukasiewicz paths



Nodes in red = n^{th} first nodes.

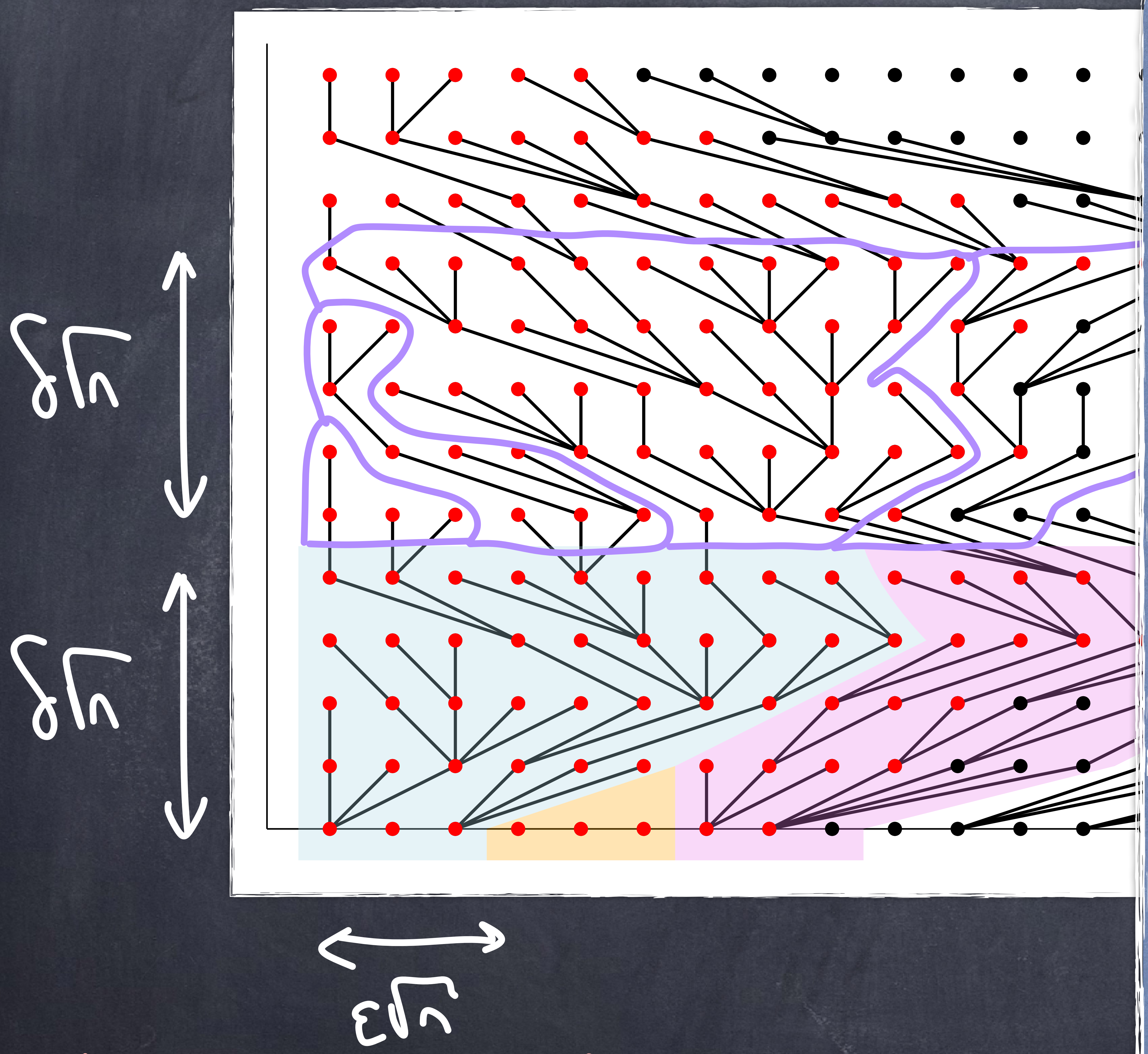
Ideas of the proof: (1) Convergence of the Łukasiewicz path



Nodes in red = n^{th} first nodes.

Main condition: $\frac{1}{n}\sum_{i=1}^{n}\sigma^2_{H(i)} \longrightarrow \sigma^2$

Ideas of the proof: ① Convergence of the Łukasiewicz path

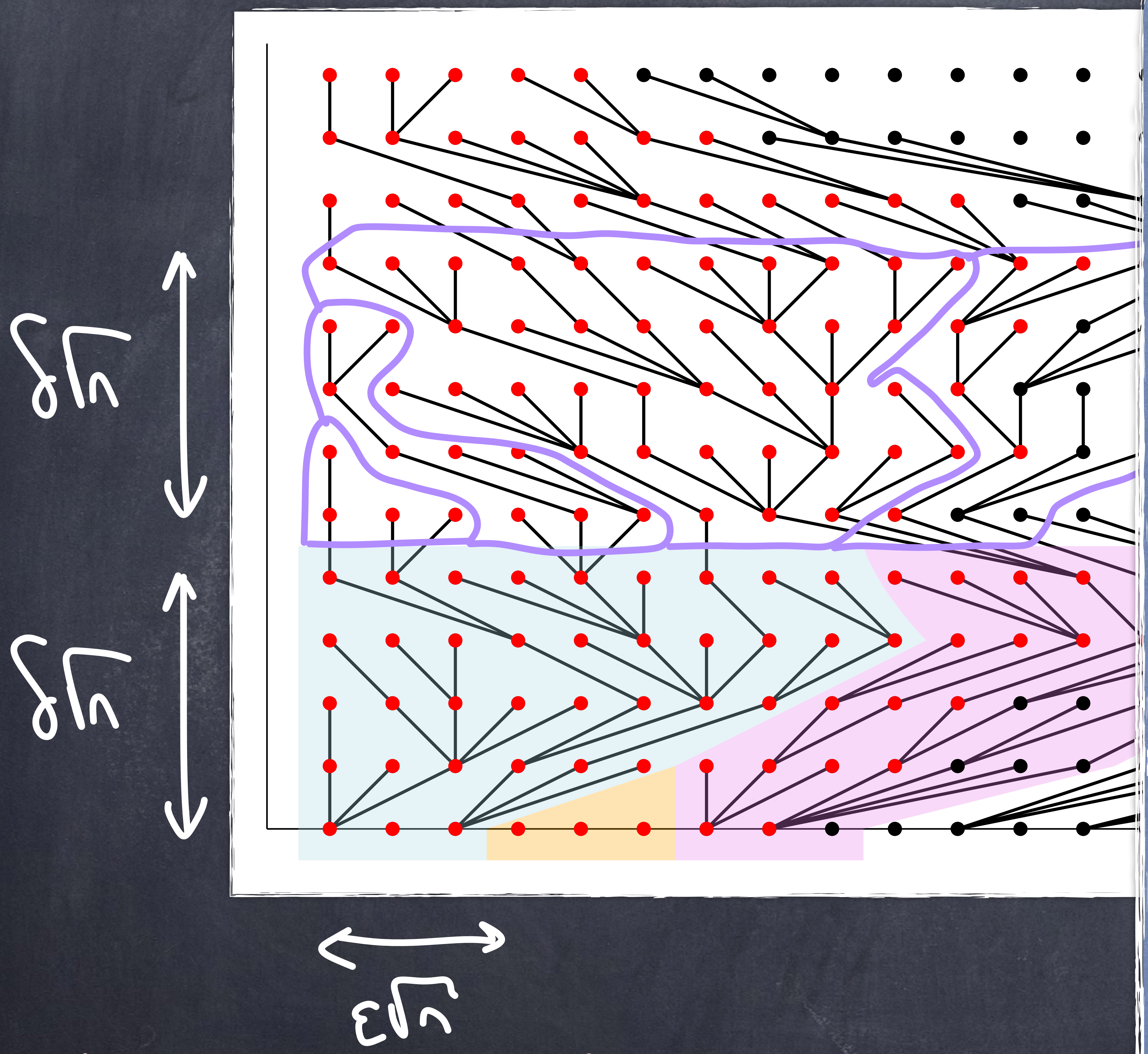


Nodes in red = n^{th} first nodes.

Main condition: $\frac{1}{n}\sum_{i=1}^{n}\sigma^2_{\mu(i)} \longrightarrow \sigma^2$

* Given $(\mu_i)_{i\geq 1}$ the blocks are iid.

Ideas of the proof: ① Convergence of the Lukasiewicz path



Nodes in red = n^{th} first nodes.

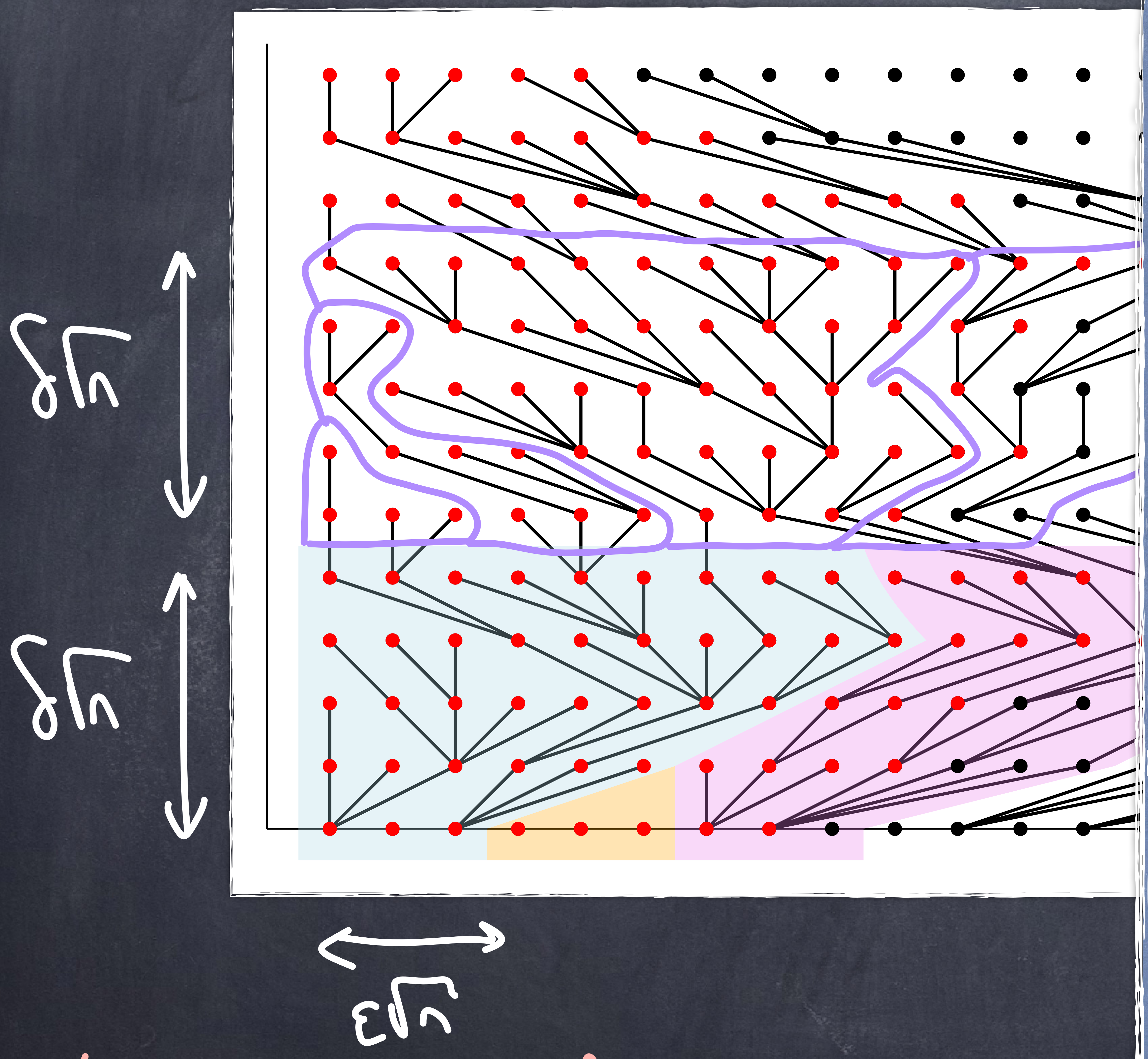
Main condition: $\frac{1}{n}\sum_{i=1}^{n} \sigma^2_{H(i)} \longrightarrow \sigma^2$

* Given $(\mu_i)_{i\geq 1}$ the blocks are iid.

* Whp as $n\uparrow\infty$, for all blocks B,

$$\sum_{v\in B} \sigma^2_{h(v)} \approx \sigma^2 |B| \qquad (*)$$

Ideas of the proof: ① Convergence of the Łukasiewicz path



Nodes in red = n^{th} first nodes.

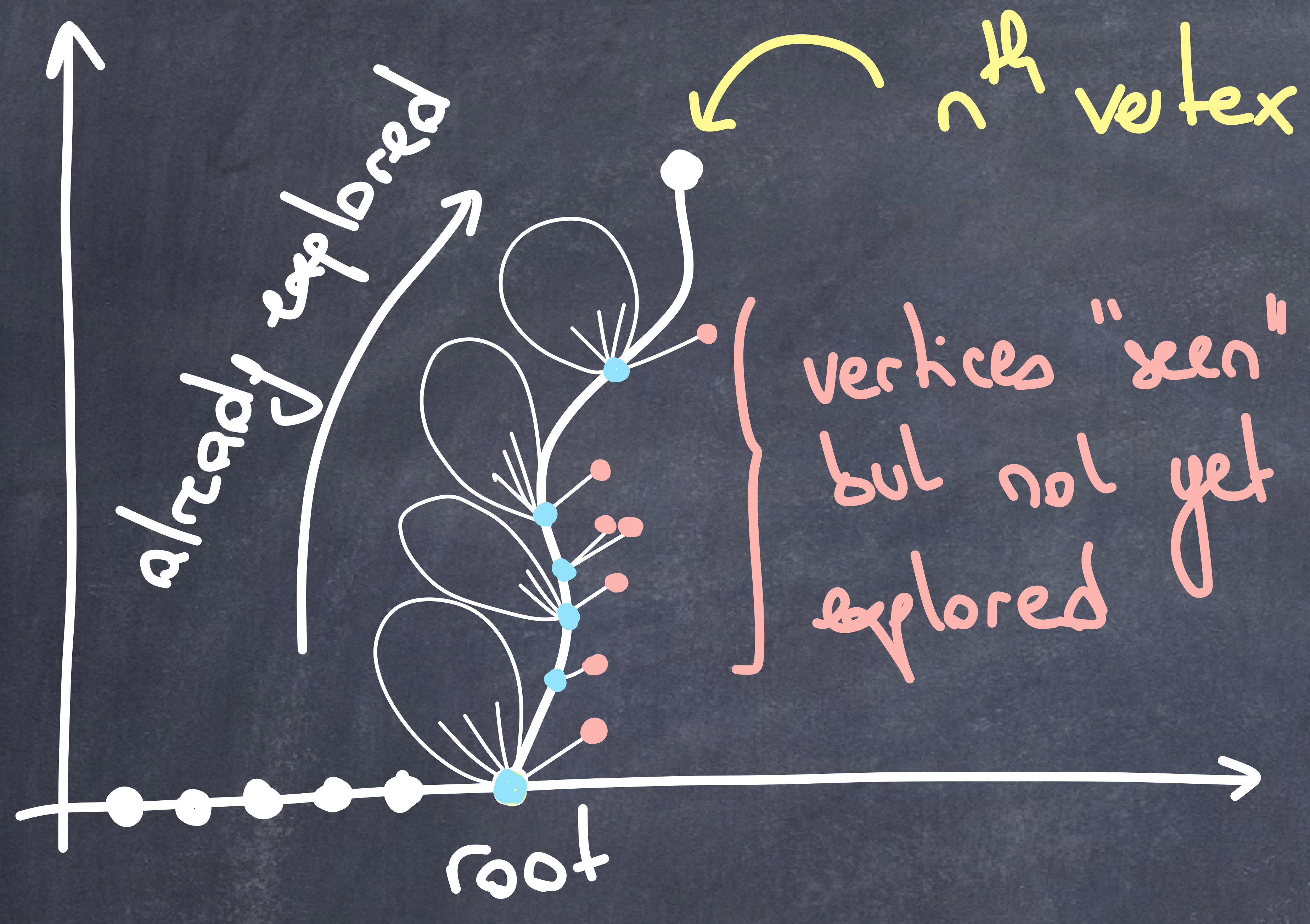
Main condition: $\frac{1}{n}\sum_{i=1}^{n}\sigma^2_{H(i)} \longrightarrow \sigma^2$

* Given $(\mu_i)_{i\geq 1}$, the blocks are iid.
* Whp as $n\uparrow\infty$, for all blocks B,

$$\sum_{v\in B}\sigma^2_{h(v)} \approx \sigma^2|B| \qquad (*)$$

* Blocks that do not satisfy $(*)$ & blocks on the boundary are negligible

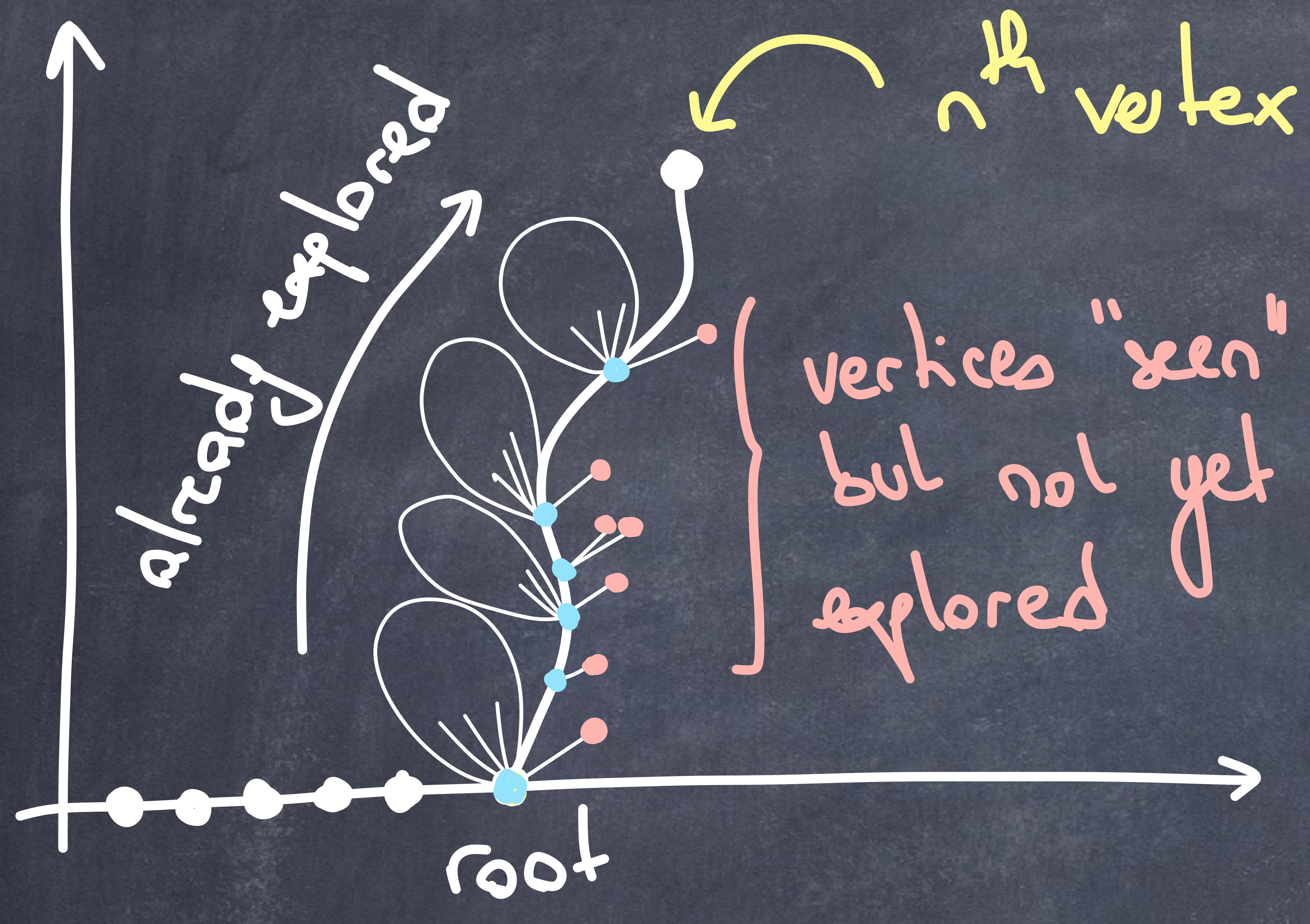
Ideas of the proof: ② Convergence of the height process



$H(n)$ = blue vertices

$L(n) - \underbrace{\min_{j \le n} L(j)}_{\#\text{ of previous trees in the forest}}$ = pink vertices.

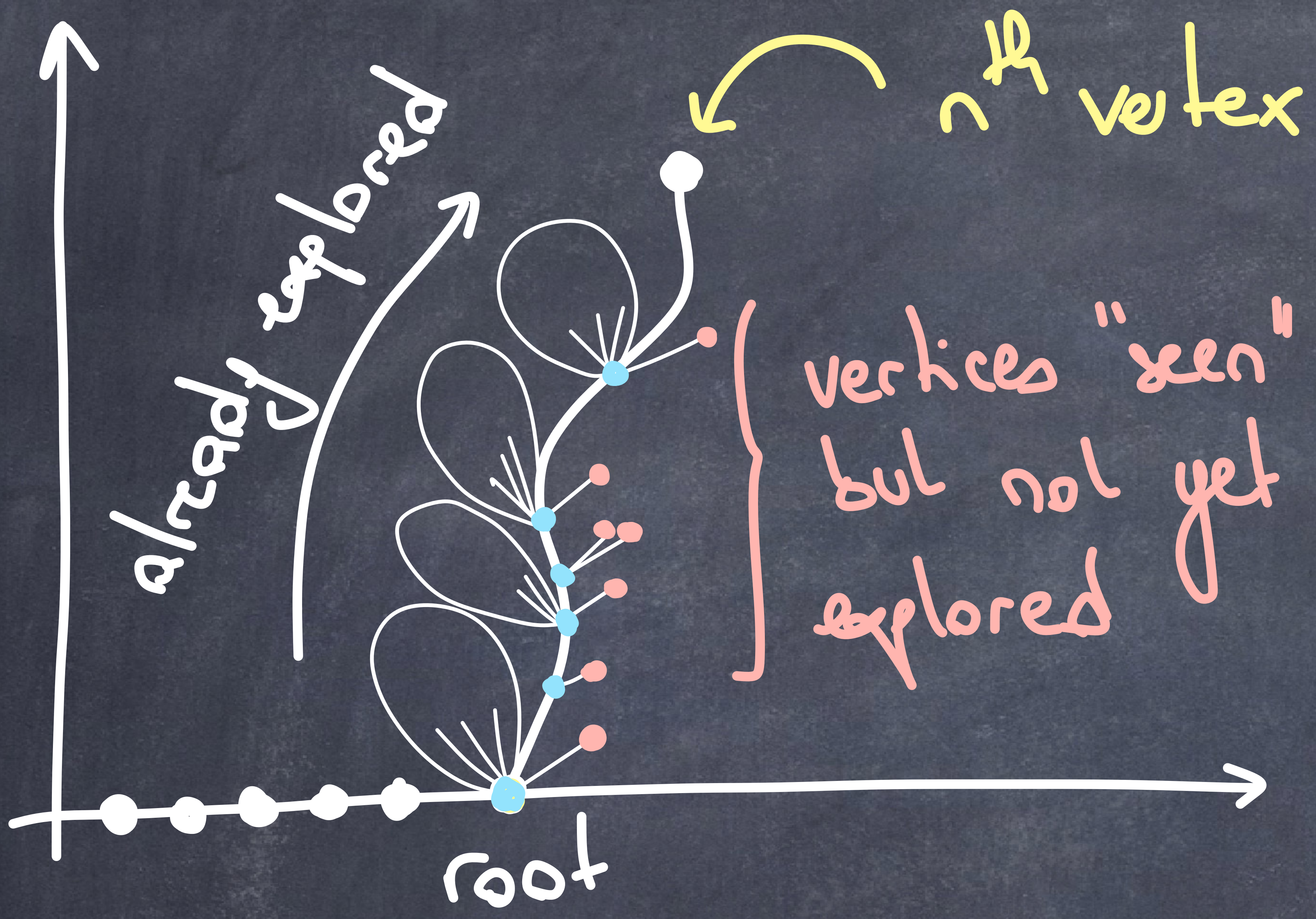
Ideas of the proof: 2 Convergence of the height process



$H(n)$ = blue vertices

$$\underbrace{L(n) - \min_{j \le n} L(j)}_{\#\text{ of previous trees in the forest}} = \text{pink vertices.}$$

Ideas of the proof: ② Convergence of the height process



$H(n)$ = blue vertices

$$L(n) - \underbrace{\min_{j \le n} L(j)}_{\#\text{ of previous trees in the forest}} = \text{pink vertices.}$$

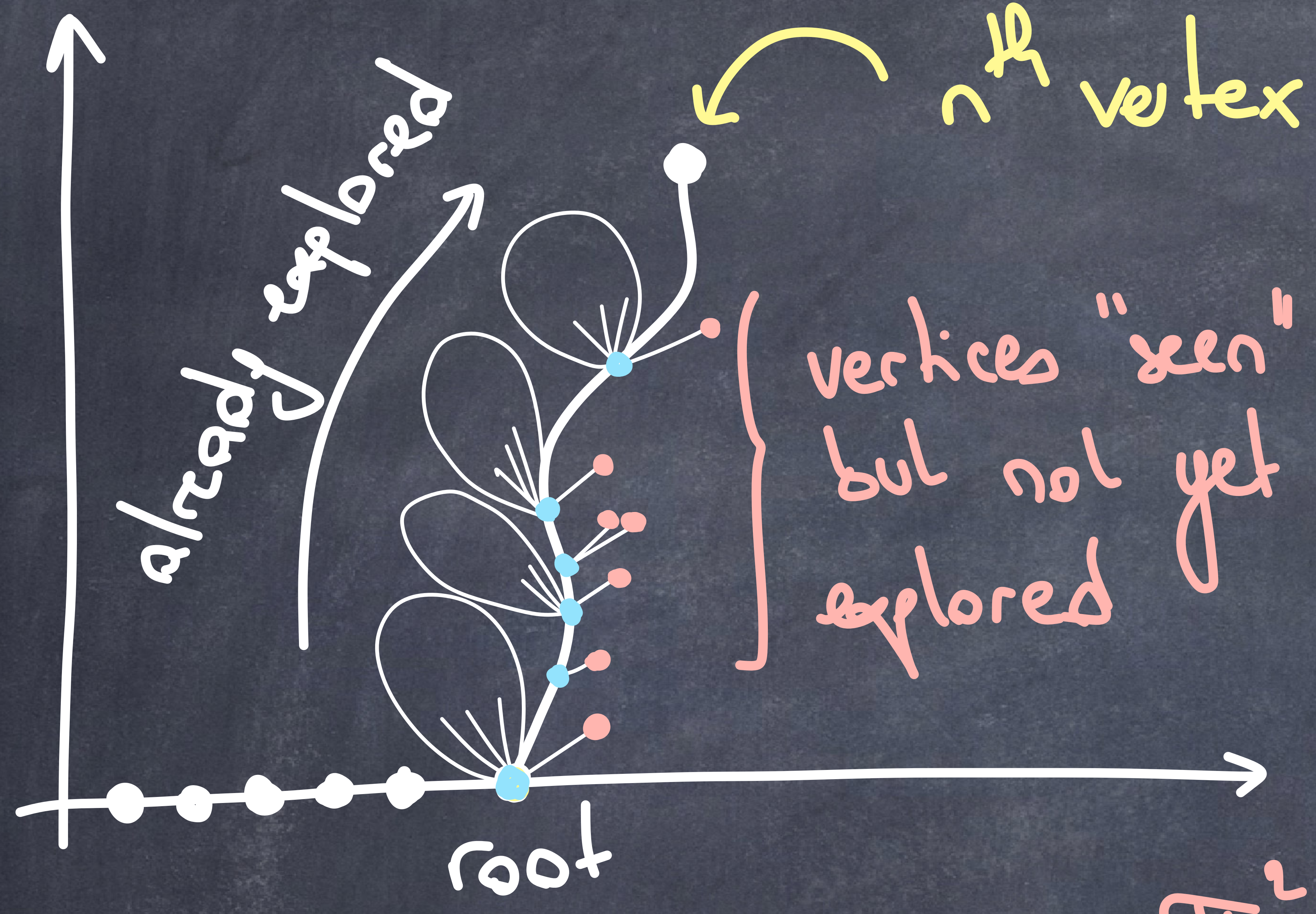
$\xi \sim$ size biased μ_t

means $\mathbb{P}(\xi = i) = i\,\mu_t(i)$

$\rightsquigarrow \mathbb{E}\xi = \sum_{i \geq 0} i^2 \mu_t(i)$

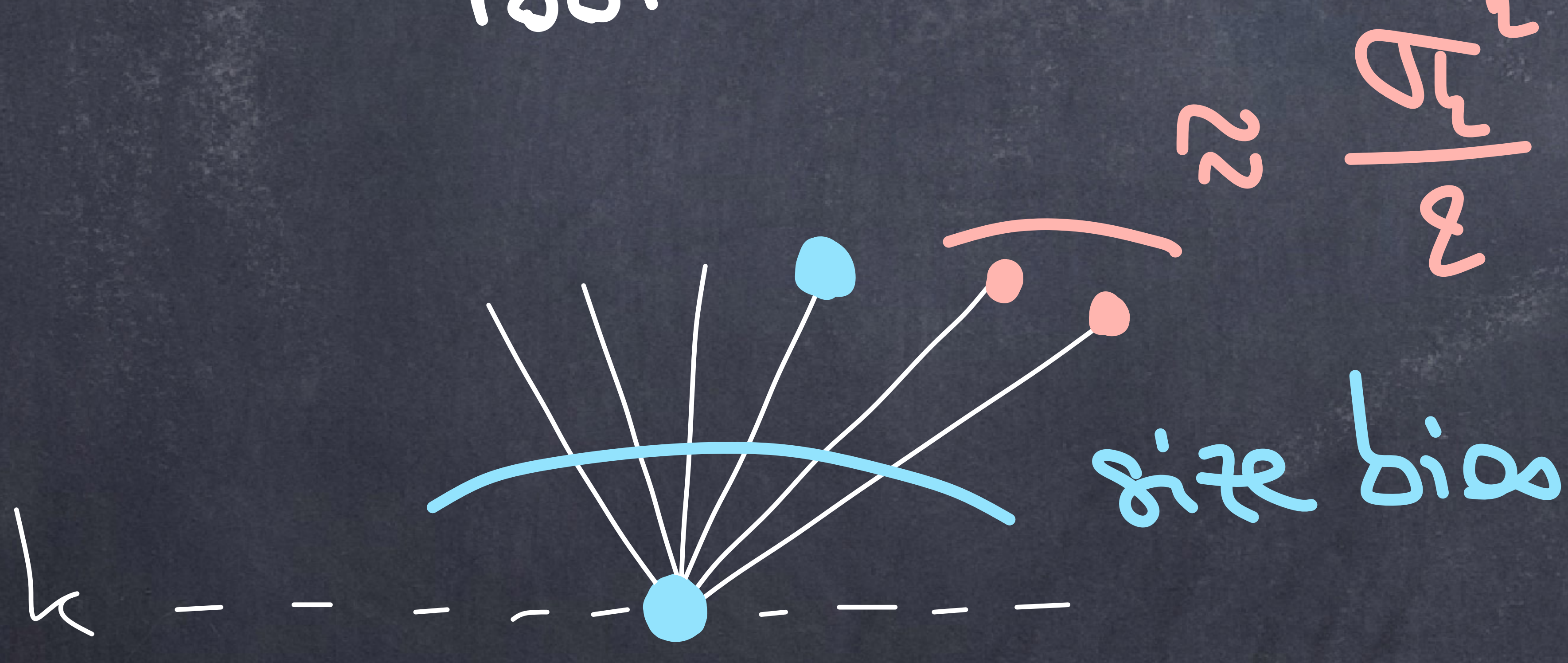
$= 1 + \sigma_t^2$

Ideas of the proof: ② Convergence of the height process



$H(n) =$ blue vertices

$$\underbrace{L(n) - \min_{j \leq n} L(j)}_{\#\text{ of previous trees in the forest}} = \text{pink vertices.}$$



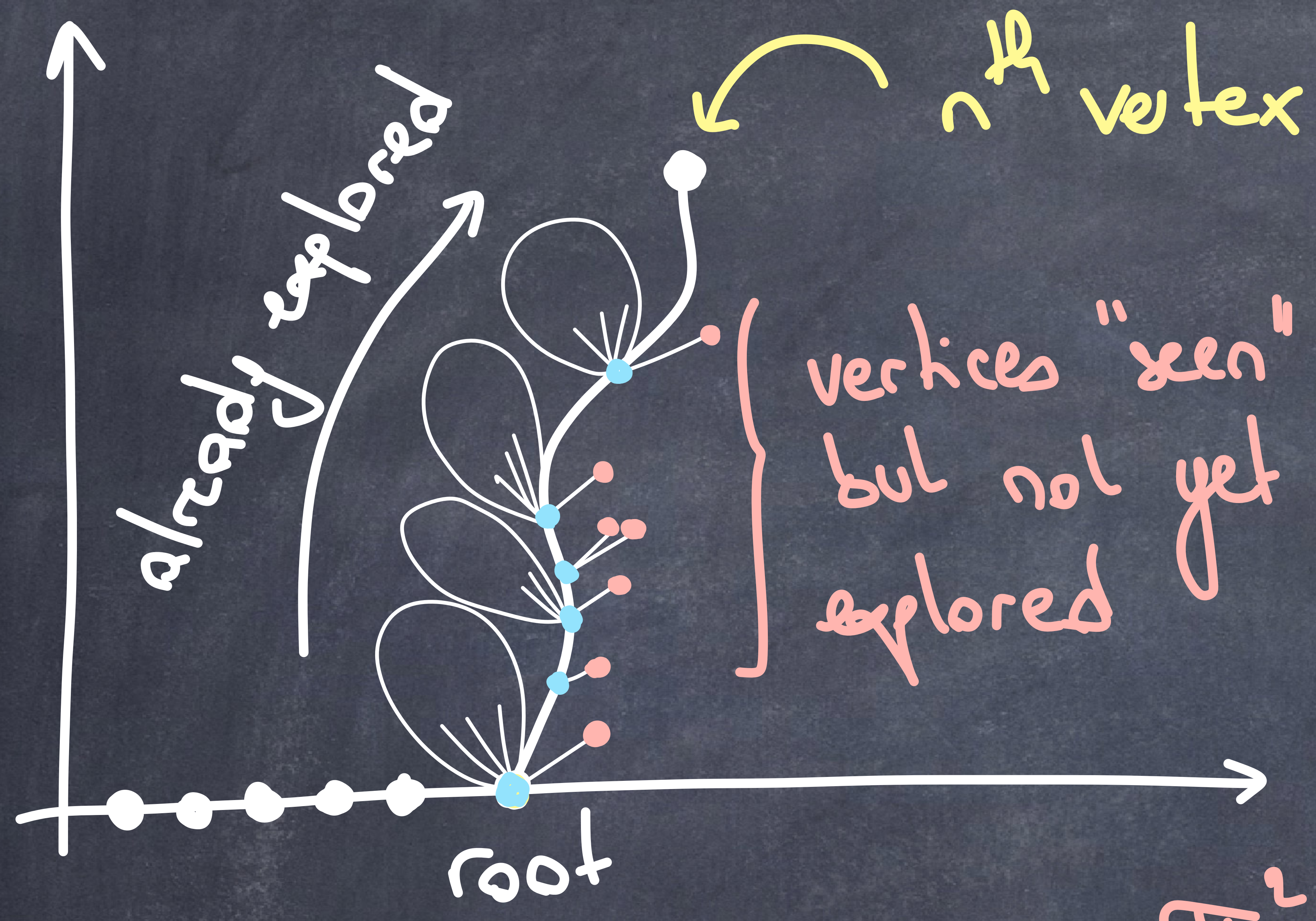
$\xi \sim$ size biased μ_ξ

means $\mathbb{P}(\xi = i) = i\mu_\xi(i)$

$\rightsquigarrow \mathbb{E}\xi = \sum_{i \geq 0} i^2 \mu_\xi(i)$

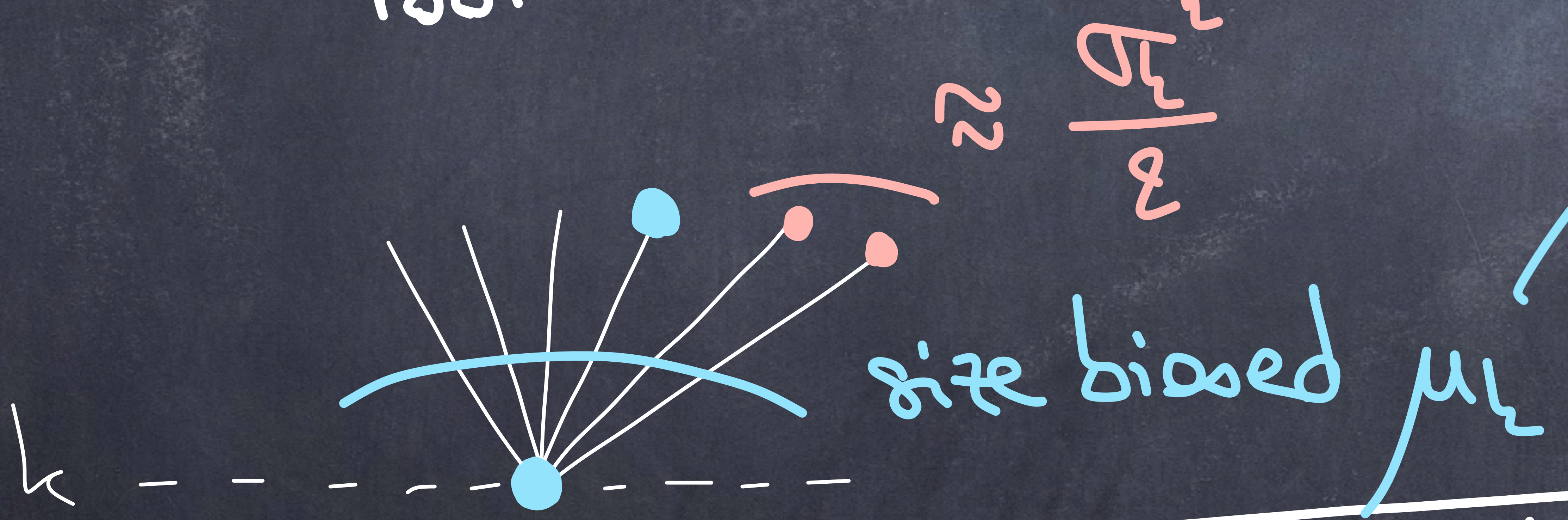
$= 1 + \sigma_\xi^2$

Ideas of the proof: ② Convergence of the height process



$H(n)$ = blue vertices

$\underbrace{L(n) - \min_{j \le n} L(j)}_{\#\text{ of previous trees in the forest}}$ = pink vertices.



$\xi \sim$ size biased μ_t

means $\mathbb{P}(\xi = i) = i\,\mu_t(i)$

$\rightsquigarrow \mathbb{E}\xi = \sum_{i \ge 0} i^2 \mu_t(i)$

$= 1 + \sigma_t^2$

$$\Longrightarrow \boxed{L(n) - \min_{j \le n} L(j) \approx \frac{1}{2} \sum_{t=0}^{H(n)} \sigma_t^2}$$

Ideas of the proof: ② Convergence of the height process

$$\boxed{L(n) - \min_{j \leq n} L(j) \approx \frac{1}{2} \sum_{k=0}^{H(n)} \sigma_k^2} \overset{LLN}{\approx} \frac{\sigma^2}{2} H(n)$$

Recall that $\left(\frac{L(\lfloor nt \rfloor)}{\sigma \sqrt{n}} \right)_{[0,1]} \longrightarrow$ Bn

Ideas of the proof: ② Convergence of the height process

$$\boxed{L(n) - \min_{j \le n} L(j) \approx \frac{1}{2} \sum_{i=0}^{H(n)} \sigma_i^2} \overset{\text{LLN}}{\approx} \frac{\sigma^2}{2} H(n)$$

Recall that $\left(\frac{L(\lfloor nt \rfloor)}{\sigma \sqrt{n}} \right)_{[0,1]} \longrightarrow Bt$

This implies $\left(\frac{L(\lfloor nt \rfloor) - \min_{j \le \lfloor nt \rfloor} L(j)}{\sigma \sqrt{nt}} \right)_{[0,1]} \longrightarrow |Bt|$

$$\Longrightarrow \left(\frac{H(\lfloor nt \rfloor)}{2\sqrt{nt}/\sigma} \right)_{[0,1]} \longrightarrow |Bt|$$

Ideas of the proof: ③ Deduce convergence of $\mathscr{L}_n$

Same as for the classical case.

⚠ it would be difficult to condition the tree on having size exactly n. (Aldous 93, Marchal-Marckert-Mokkadem 03)

Ideas of the proof: ③ Deduce convergence of $\mathcal{L}_n$

Same as for the classical case.

⚠ it would be difficult to condition the tree on having size <u>exactly</u> n. (Aldous 93, Marchal-Mokkadem 03)

<u>Open problems</u>: * $\sigma^2 = \infty$?

* non-strictly critical case ?

Ideas of the proof: (3) Deduce convergence of $\mathcal{T}_n$

Same as for the classical case.

⚠ it would be difficult to condition the tree on having size <u>exactly</u> n. (Aldous 93, Marchal-Miermont 03)

<u>Open problems</u>:
* $\sigma^2 = \infty$?
* non-strictly critical case ?

Thanks ! :)