# The # of possible evolution histories

Motivation:
- Genome = $L_0$ sites
- Mutation = single site mutations
- ancestral type = 0 , mutant type = 1
- A Genome = $\sigma \in \{0,1\}^{L_0}$ → Kauffman-Levin : N peptides constrained to use 2 amino-acids (Leucine Alanine)

Each genetic type $\sigma \longleftrightarrow$ a fitness value $x_\sigma$

Regime: low mutation , strong selection
Only mutations which increase fitness fixate.

→ Only possible evolutionary path are s.t. fitness ↗.

Kauffman Levin : One can think of the adaptive process as a continuous time, discrete state Markov process, in which the entire pop. is resident at one state and then jumps with fixed proba. to each of its 1-step mutant fitter variants

optimum at distance $L \leq L_0$ : initial state $\underbrace{(0, \dots, 0)}$ optimum $\underbrace{(1, \dots, 1)}_{\text{length } L}$

Question: how to pick the $x_\sigma$? (The fitness landscape)

smooth landscape $x_\sigma = \alpha |\sigma| = \alpha \,\#$ mutations.

The "house of cards" model $(x_\sigma)_\sigma$ are iid $\sim$ atomless law (say unif. on $[0,1]$)

LK (or NK) model $\forall i \leq L$ $\vec{\nu}_i = \{i, \nu_{i1}, \dots, \nu_{iK}\}$ : interaction neighborhood

$x_\sigma = \sum_{i=1}^{L} f_i(\{\sigma_j\} \; j \in \vec{\nu}_i)$ where the $f_i$ are iid as $\sigma_{\vec{\nu}_j}$ takes its $2^{K+1}$ possible values

$K = L-1 \rightarrow$ HoC.

## Neutral model viability with proba $p$ (iid Bernoulli)

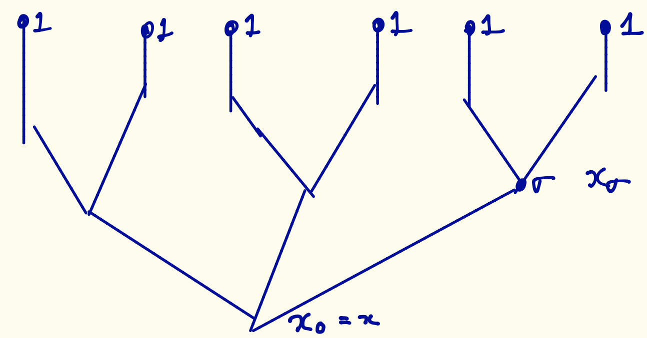
$(x_\sigma, \sigma \in \{0,1\}^L)$ iid unif on $[0,1]$ except: $x_{\sigma_0} = x$ $\quad x_{\sigma_L} = 1$

$\Theta = \#\{$ open paths $\sigma_0 \to \sigma_L$ $\forall i$ go from $\sigma_i$ to $\sigma_{i+1}$ by flipping a $0$ into a $1$ $\}$

**Toy model**: Tree $\emptyset$ has $L$ offsprings

$1^{st}$ gen each indiv has $L-1$ offsprings

$k^{th}$ gen.: each indiv has $L-k$ offsprings



## Main results

$\mathbb{P}_x(\cdot)$ law of global R.V. $\Theta$ when $x_{\sigma_0} = x$

$\mathbb{P}^*$ ———————— $x_{\sigma_0} \sim \mathcal{U}_{[0,1]}$

$$\mathbb{E}_x(\Theta) = L! \cdot \frac{(1-x)^{L-1}}{(L-1)!} = L(1-x)^{L-1}$$ Tree and cube

If $x=0$ $\mathbb{E}_0(\Theta) = L$ $\mathbb{E}^*(\Theta) = 1$ Tree and cube

Proposition $\lambda > 0$ $\lim_L \mathbb{E}_{\lambda/L}(\Theta/L) = e^{-\lambda}$

$\lim_L \mathrm{Var}_{\lambda/L}(\Theta/L) = e^{-2\lambda}$ Tree

$\lim_L \mathbb{E}_{\lambda/L}[\Theta/L] = e^{-\lambda}$, $\lim_L \mathrm{Var}_{\frac{\lambda}{L}}[\Theta/L] = 3e^{-2\lambda}$ cube

If $x = \frac{\lambda}{L}$ L large Then $\mathrm{Var}(\theta) \sim E(\theta)^2$ So if $x = O(\frac{1}{L})$ $\Theta \simeq E[\Theta]$

if $x \sim \mathcal{U}_{[0,1]}$ $P(x = O(1/L)) = O(1/L)$ giving $O(L)$ paths

$\to \mathbb{E}(\theta) = 1$ but $\mathrm{Var}(\theta) \sim L$ !

Thm Tree: If $x = \frac{\lambda}{L}$ $\frac{\theta}{L} \xrightarrow{\mathcal{L}} e^{-\lambda}\mathcal{E}$ where $\mathcal{E} \sim \exp(1)$ R.V.

cube If $x = \frac{\lambda}{L}$ $\frac{\Theta}{L} \xrightarrow{\mathcal{L}} e^{-\lambda}\mathcal{E} \times \mathcal{E}'$ where $\mathcal{E}, \mathcal{E}' = 2$ indep $\exp(1)$ R.V.

Thm Tree: If $x = \frac{\ln L + y}{L}$ $\lim_{L} \mathbb{E}_x[\Theta] = e^{-y}$

$\lim_{L} \mathrm{Var}_x(\theta) = e^{-2y} + e^{-y}$

$$\mathbb{P}^*[\theta \geqslant 1] \sim \frac{\ln L}{L}$$

$\mathbb{E}$ and Var of $\Theta$ dominated by $x \sim 1/L$

$\mathbb{P}_x(\theta \geqslant 1)$ dominated by $x \sim \frac{\ln L}{L} + O\left(\frac{1}{L}\right)$

Discussion :

Novak & Krug : n-Tree with $n(h)$: $\mathbb{P}(\overset{\swarrow \text{\# open path up to } h.}{\textcircled{H}} \geq 1) \sim \frac{n(h)^h}{h!}$ if $n$ is constant or sublinear

If $n(h) = \alpha h$ $\exists \alpha_c \in [\frac{1}{e}, 1]$ s.t. $\mathbb{P}(\textcircled{H} \geq 1) = \begin{cases} \to 0 & \alpha < \alpha_c \\ \to > 0 & \alpha > \alpha_c \end{cases}$

Roberts & Zhao Zhao $\alpha_c = \frac{1}{e}$ and at $\alpha_c \to 1$.

Hegarty - Martinson $\left.\begin{array}{l} \lim_{L} \mathbb{P}_{\frac{\ln L}{L} - \varepsilon_L}(\Theta \geq 1) = 1 \\ \phantom{\lim_{L} \mathbb{P}_{\frac{\ln L}{L}}} + \varepsilon_L \phantom{(\Theta \geq 1)} = 0 \end{array}\right\} \Rightarrow \mathbb{P}^*(\Theta > 0) \sim \frac{\ln L}{L}$

## Proof idea $\quad x = \frac{y}{L} \quad y \geq 0$ fixed

Natural idea: $G(x,\lambda,L) := \mathbb{E}_x[e^{-\lambda \theta}]$

$$G(x,\lambda,1) = e^{-\lambda}$$

$$G(\lambda,x,L) = \prod_{i=1}^{L} \mathbb{E}\left[e^{-\lambda \,\#\text{ open from } i} \mathbb{1}_{x_i > x} + \mathbb{1}_{x_i < x}\right]$$

$$= \left[x + \int_x^1 dz\, G(\lambda, z, L-1)\right]^L$$

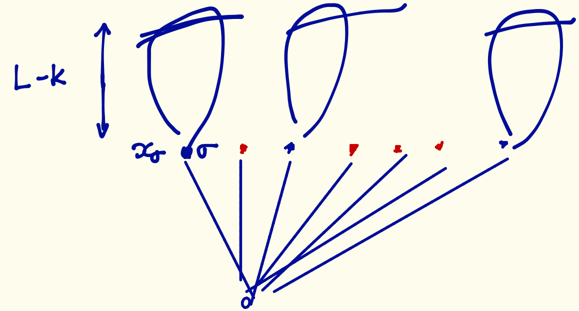
pb: The # of levels and the size of levels change at each step.

# Tree

$\Rightarrow$ The value of $\frac{\Theta}{L}$ is decided at early stages of branching process

$\mathcal{F}_k$ = info of $k$ $1^{st}$ levels

$$\Theta_k = \mathbb{E}[\Theta \mid \mathcal{F}_k] = \sum_{|\sigma|=k} \mathbb{1}_{\sigma_{open}} (L-k)(1-x_\sigma)^{L-k-1}$$




as $k \nearrow \infty$ LLN $\Theta_k$ becomes good approx of $\Theta$

$1^{st}$ $L \to \infty$, then $k \to \infty$

2 Steps: 1) Show $\lim_L \mathbb{P}_{y/L}\left(\frac{\Theta}{L} \leq \gamma\right) = \lim_k \lim_L \mathbb{P}_{y/L}\left(\frac{\Theta_k}{L} \leq \gamma\right)$

2) Then Show dist $\frac{\Theta_k}{L} \rightsquigarrow$ exp law.

Step 1:

upper bound $\mathbb{P}\left(\frac{\theta}{L} \leqslant \xi \mid \mathcal{F}_k\right) \leqslant \mathbb{1}\left\{\frac{\theta_k}{L} \leqslant \xi + \delta\right. \left|\mathcal{F}_k\right] + \mathbb{P}\left[\frac{|\theta - \theta_k|}{L} \geqslant \delta \mid \mathcal{F}_k\right]$

Chebyshev $\mathbb{P}\left(\frac{|\theta - \theta_k|}{L} \geqslant \delta \mid \mathcal{F}_k\right) \leqslant \frac{\mathrm{Var}(\theta \mid \mathcal{F}_k)}{L^2 \delta^2}$

$$\Rightarrow \mathbb{P}_{y/L}\left(\frac{\theta}{L} \leqslant \xi\right) \leqslant \mathbb{P}_{y/L}\left(\frac{\theta_k}{L} \leqslant \xi + \delta\right) + \frac{\mathbb{E}_{y/L}\left[\mathrm{Var}(\theta / \mathcal{F}_k)\right]}{L^2 \delta^2}$$

Need: $\lim_k \limsup_L \frac{1}{L^2} \mathbb{E}_{y/L}\left[\mathrm{Var}(\theta \mid \mathcal{F}_k)\right] = 0$

lower bound same argument

Step 2 Law of $\theta_k$ $$G_k(\lambda, x, L) = \mathbb{E}_x\left[e^{-\lambda \theta_k}\right]$$

$$G_0(\lambda, x, L) = \exp\left(-\lambda L (1-x)^{L-1}\right]$$